

Joys of darkness

Planetaria are getting ever better at capturing the glories and the challenges of the night sky. They are underused not only in that astronomical context, but for other sciences too.

Of all people, New Yorkers face the biggest obstacles to experiencing the magic of a moonless, starry night. The absolute quiet, the spangled black panorama, the misty billions of stars in the Milky Way, the sense of majestic timelessness. No wonder, having lost all this, New Yorkers top the league tables for rancour. But they have been partially compensated: since February, they have been able to enjoy the Hayden Planetarium at the American Museum of Natural History.

For this technological achievement and for the great spectacle it provides, all credit to the engineers of the Carl Zeiss planetarium division in Jena, eastern Germany; to the Hubble Space Telescope, which is the source of many of the planetarium's best images; and to countless developments over the past decade or two in the understanding of the structure and content of our stellar neighbourhood, of our Galaxy as a whole, and of the largest scales of the detectable Universe.

The Hayden's technology is worth highlighting. The "Universarium" projector, like conventional planetarium projectors, uses an arc lamp to give appropriate colour balances and star masks by which the stellar images are projected onto the dome. One innovation is to use fibre optics to concentrate the light onto the holes in the masks and thereby allow a hundredfold increase in brilliance. Another is to draw on stellar position data from an unprecedentedly large database compiled by the US space agency NASA, allowing stars to be projected that are below most people's thresholds of visible detection. (Visitors can bring binoculars to see for themselves.) Then there are the more conventional facilities — separate projectors for more mobile sky objects such as the planets and the Moon, multiple projectors for conventional images spanning the dome's surface, and computer-simulation projection.

Fulfilment and frustration

The half-hour show *Passport to the Universe* is routinely sold out at the Hayden several times a day. Anyone with any knowledge of astronomy — even an amateur acquaintance with the night sky — is certain to find it fulfilling and frustrating. Fulfilling not only because of the sheer quality of the images, but also because nowhere else will they have seen such a fine and up-to-date visualization of the distribution of superclusters of galaxies that makes up the largest-scale structure of the Universe, or such a superb rendering of an imaginary trip through the Orion nebula, a 'nearby' gas cloud in which stars are being formed. Frustrating, however, because these are merely glimpsed and yet contain glories and mysteries enough in themselves to sustain more extensive display — the dynamics of motions of groups of galaxies, for example, and the evolution of the Universe's structure. Included in the show's portrait of the Orion nebula, but unremarked, is a protostar complete with accreting gas and jets of ejected material — another fine feature to focus on, with a possible speculative extension into the growing numbers of stars around which have been detected preplanetary clouds, protoplanetary disks, brown dwarfs and planets.

Such phenomena are at the forefront of current research. But it

would be regrettable if, as the imaging improves, newly discovered phenomena crowded out the traditional portrayal and explanation of well-known features of our stellar neighbourhood: long-understood objects visible as mere dots with the naked eye, such as the zoo of variable stars, both single (such as γ Cassiopeia, a white-hot star that discards material in occasional bursts) and multiple (including distorted interacting stars as they orbit their common centre of mass); and the variety of beautiful star clusters, 'open' (such as the Pleiades) and globular (ω Centauri).

Opportunities

Given the popularity of the Hayden shows, it is clear that there is an appetite for far more to be shown. And with at least 1,500 planetaria worldwide, albeit less sophisticated than that in New York, it is also clear that the opportunities for better exploitation of such facilities are global and growing.

Many planetaria are to be found at science centres, yet seem all too often to be considered as purely astronomical show-places. Much of the technical effort is devoted to precise reproduction of stellar and planetary motions, and to perspectives from different Galactic locations and eras. But planetaria possess more conventional projection facilities that will surely blossom if allied with two other technologies: supercomputing simulation and computer-based animation. Might most planetaria become 'scientaria'?

Visit any supercomputing facility and you will probably encounter simulations not only of astronomical processes but also of molecular interactions, electronic materials and complex fluid flows. Enclose the observer in a planetarium hemisphere and project these with due imagination, and you could quite quickly compel their attention to the complex beauty of the processes at work, thereby, incidentally, unlocking some of the black boxes of everyday technology. Turn to the animators, and imaginary but scientifically insightful voyages through cells become a possibility. Given very recent progress, the structure and function of key cell constituents such as ribosomes and membrane channels can now be speculatively but credibly visualized. What better way to put into biological context the progress of the Human Genome Project than by portraying the chain of action from genome to individual gene expressions and their macroscopic outcomes, and how these are influenced by environmental signals?

Whether the current planetarium suppliers — Zeiss, Spitz and the like — or the growing numbers of more conventional 'immersive theatre' companies can profitably fulfil this opportunity remains to be seen. But millions of people every year would benefit from the results. And of course, all of this is but a step towards virtual reality, in which any observer can happily immerse themselves in his or her own simulation or database, and which supercomputer centres are already exploiting.

Immersive visualization is not the same as understanding, but, as New Yorkers might well agree, it's a great step along the way, and far closer to the magic of the real world, night and day, than the Discovery channel. ■





Up in smoke
Forest fire research
falls victim to the
flames
p5



Dope tests
California backs
research into
cannabis
p6



Global warming
Report offers
carbon dioxide
a reprieve
p7



Head start
More Neanderthal
bones unearthed
in Germany
p9

Britain may boost protection of researchers from intimidation...

Natasha Loder, London

Prompted by the increasingly aggressive tactics of animal-rights activists, Britain's Home Secretary, Jack Straw, promised last week to consider changes in the law that would give the police more powers to stop individuals organizing campaigns of intimidation.

His comments came after five firebomb attacks on cars belonging to workers at Huntingdon Life Sciences (HLS), a contract company that uses animals mainly to test medicines for safety for the pharmaceuticals industry, and agricultural chemicals for toxicity.

No group has claimed responsibility for the attacks. But it is widely assumed that they were the work of animal-rights activists. "Most of the hardcore activists are central to the campaign against Huntingdon," says Mark Matfield of the Research Defence Society, a body set up to protect the use of animals in research.

Since 1997, protest campaigns have closed three animal breeding centres. Britain's only supplier of primates for research, Shamrock Farm, shut earlier this year (see *Nature* 404, 215; 2000). It had been the site of confrontations between masked protesters and the police, and its managing director suffered a series of attacks.

In 1997, Consort Kennels of Ross-on-Wye, which bred beagles for use in research,



Terror campaign: cars have been firebombed in an attempt to close down Huntingdon Life Sciences.

closed after a campaign of attacks on the homes of its directors. And Hillgrove Farm near Oxford, which supplied cats for experiments, closed in 1999 after letter-bombs were sent to the farm's owner and two employees.

Andrew Gay of HLS says the law has failed to protect his staff against death threats and abusive letters. Forty staff are shareholders in the company, and their names and addresses have been released, as required by British law.

Earlier this year, 11% of HLS shares were dumped by city fund managers when directors' names were publicized, and a campaign of abuse, telephone calls, letters and bomb

threats followed. Gay says that, despite the campaign, orders for experimental animals have grown by 40% in the past two years.

Home Office minister Mike O'Brien visited HLS and met staff in July, and has launched a series of consultations with police, scientists and animal-welfare organizations. The government has previously looked at whether tighter controls are needed on animal experimentation. It now says it is determined to curb the intimidation of scientific and laboratory staff.

Many scientists believe that animal activists are dealt with too lightly by the courts because they claim to be acting out of idealism rather than self-interest. The Home Office plans to discuss the matter with the department that manages UK courts. It will also liaise with the Association of Chief Police Officers to discuss the policing of animal-rights demonstrations and the effective use of existing laws.

Over the next three months the Home Office will consider extending the law in the area covering incitement to intimidate. But a heavy parliamentary timetable means that any new legislation would have to wait until after the next general election, which may take place next year.

Cambridgeshire police — which has a team of 40 officers investigating animal-rights activists in the region — claims that loopholes in current harassment legislation have prevented arrests.

...as German activists wage propaganda war

Quirin Schiermeier, Munich

Animal-rights activists in Munich have been accused of using confidential information leaked from a Bavarian ethics commission as the basis of a campaign against experiments on Java macaques.

Konrad Messmer, director of the institute of surgery research at the Ludwig-Maximilian University in Munich, also claims that a leak formed the basis of a deliberate media disinformation campaign. The activists say the information

was obtained via legal channels.

Several newspapers and radio stations in Bavaria last week reported that a researcher at the institute intended to amputate the arms and legs of the macaques for xenotransplantation studies. The researcher and his family were inundated with aggressive telephone calls from animal-rights activists and animal lovers.

But the experiments, which were approved by the Bavarian authorities in July, merely involve

the microscopic analysis of the macaques' thigh muscles under anaesthetic. The research is designed to investigate the mechanisms of rejection reactions in human organ transplants. Messmer says the experiments are not related to xenotransplantation, nor do they involve amputation.

In response to the accusations, research lobbyists in Munich are now considering taking legal action against unknown members of the ethics commission.



Ngubane (left) and Northern Cape premier Manne Dipico with replicas of the telescope.

Construction starts on largest telescope in the south

Michael Cherry, Cape Town

Last Friday, work began outside the town of Sutherland, about 400 kilometres north-east of Cape Town, on building the largest telescope in the southern hemisphere.

The Southern African Large Telescope (SALT), due to be completed by the end of 2004, is similar in design to the Hobby-Eberly Telescope at Fort Davis in Texas, and will give astronomers an equivalent facility for the southern hemisphere.

SALT can access 70% of the sky observable at Sutherland, and will gather 20 times more light than any other African telescope. This allows observation of many important objects from the southern hemisphere, such as the centre of the Milky Way and the two nearest galaxies, the Magellanic Clouds.

The cost of construction and operation for the first ten years will be around R200 million (US\$30 million), a target that project scientist David Buckley says is just short of being met. Of this, 38% will come from the South African government, through the National Research Foundation.

The rest of the money is coming from universities in the United States, New Zealand and Germany, and from the Polish Academy of Sciences, which will receive a *pro rata* allocation of viewing time. Latest to join has been a consortium of UK institutions.

"The telescope will provide a focus for the development of basic sciences on the African continent," said Ben Ngubane, minister of arts, culture, science and technology at the opening ceremony.

A novel aspect of the project is a collateral benefits plan, which aims to ensure that SALT benefits South African society. For example, as many components as possible will be made locally.

Japan's desire to lead in IT fuels latest budget requests

David Cyranoski, Tokyo

Information technology (IT) fever has hit Japan's ministerial budget requests for next year, which were finalized last week. The government's mood was reflected in a promise by prime minister Yoshiro Mori's IT strategy panel that Japan will overtake the United States in IT in five years.

The ministries have responded accordingly. Increased funding is also slated for new programmes in nanotechnology, immunology and bioinformatics (see table). But brain, nuclear and space research all face decreased budgets within the Science and Technology Agency (STA).

An estimated ¥1 trillion (US\$9.4 million) will be earmarked for IT in the new budget, which is likely to be formally approved, after slight modification by the government, later this year. It includes ¥364 billion for the new National Land and Transportation Ministry and covers items such as the National Police Agency's cyberterrorism policy.

The IT budgets of the Ministry of International Trade and Industry (MITI) and the STA would both increase significantly (see table). But there has been criticism of the ill-defined goals for the extra IT spending, especially the lack of a clear description of how the new public money will correlate with private initiatives.

The Mori panel's statement will do little more than temporarily "cheer up" the research and business community, says one senior researcher, who compares it to last year's "ridiculous" promise by the Japanese

government to spur the creation of 1,000 Japanese biotech start-up companies by 2010.

But the budget requests also include some IT-related projects based firmly in areas of Japanese strength. Planners for the STA's Nanomaterial Research Center, for example, are promoting nanoscience as the key to Japan's 'IT revolution'.

A proposed budget of ¥3.3 billion a year for five years for this area will go towards facilities in Tsukuba and for gathering researchers across Japan to work on a broad range of 'nanoprojects'. Having a centre around which networks can form is a top priority for building on research that is already strong in Japan, says the centre's Tomoji Kawai.

The MITI's Correlated Electron Research Center (CERC), a proposed eight-year, \$10 million a year project, also builds on Japanese strengths. The centre, growing out of the spin electronics division of the Joint Research Center for Atom Technology (JRCAT), will be in Tsukuba, and will continue the JRCAT's policy of encouraging industrial collaboration.

Yoshinori Tokura of Tokyo University, one of those behind CERC's creation, welcomes any boost IT might give his project. But he admits that his research—focused on organic or oxide-correlated electron systems—is far from ready for application.

The STA's Nanomaterial Research Center is a response to the US National Nanotechnology Initiative, announced in January, and consists of a \$5 billion budget (an 83% increase) for nanotechnology (see *Nature* 405, 730–732; 2000). This initiative has challenged Japan to improve the organization of its research strengths to compete globally.

A MITI official says the US initiative, and visits by US government officials to Japan's nanotech research and development sites, appears to confirm that Japan is far ahead in many fields, stimulating a desire to stay on top, and leading to this year's budget requests.

Another field in which Japanese research is strong, but fragmented, is immunology. The STA has requested ¥3.6 billion of renewable five-year funding for a centre, to be located in Tsukuba. This is partly intended to be a rallying point for a research community that some feel has not received the funding its status deserves. But planners also hope it will stimulate interaction between immunology researchers, hospitals and universities.

Bioinformatics will gain a new centre that aims to become Japan's version of the US National Center for Biotechnology Information. And proteomics will get a big increase for its nuclear magnetic resonance centre, which plans to characterize the structure and function of 3,000 proteins in the next five years.

Highlights of Japan's R&D budget requests for fiscal year 2001 in ¥ billion (US\$1=¥107)

	2001 request	% change
Ministry of International Trade and Industry (MITI)		
Total R&D request	587.9	11.2
Information technology	52.6	190
Materials Nanotechnology Program	5.0	New
Medical innovation based on genomic analysis	16.3	63
Protein Functional Analysis	5.0	270
Bioinformatics	5.0	107

Science and Technology Agency (STA)

Grants-in-aid for scientific research	152.9	7.8
Information technology	46.8	120
Nanotechnology (including new centre)	15.0 (3.3)	21 (new)
Life science total	101.1	25.4
Analysis of protein structure/function	19.0	119
Bioinformatics centre	4.5	New
Immunology centre	3.6	New
SNPs centre	2.5	35
Brain research centre	10.5	–11.3
Nuclear research	318.8	–0.2
Space research	184.8	–1.2



From trees to ash: this 'Woolsey' plot in Arizona, established in 1909, was among those destroyed.

US forest fires send ecology experiments up in smoke

Paul Smaglik

Fires that have laid waste to more than 2.4 million hectares of US forest this summer represent a double dose of ecological irony.

Not only have the fires destroyed historic research plots whose increasing concentration of trees over the past 90 years foretold this summer's disaster, but they have also consumed some newly established plots designed to investigate how to prevent such massive fires from flaring up again.

Although ecologists have not yet conducted a full survey of the damage, and the fire season is not yet over, blazes have already damaged or destroyed a third of the so-called Woolsey plots, which are among the oldest research plots in the country.

The plots, named after ecologist T. S. Woolsey, who led US forest-service workers in the early 1900s in counting, measuring and mapping each tree on the 50 plots, ranged from about 2 to 6 hectares in size.

They were surveyed every five years until the 1930s, when they were abandoned. Ecologists rediscovered them in the 1990s, but by then only about 15 remained, many of the others being lost to development.

The plots give ecologists a detailed history of the forest and its changing composition. They also document how the concentration of trees increased when the forest service decided to prevent fires in forests, rather than allow periodic ground fires to remove underbrush and dead trees.

The first time the foresters surveyed the plots there were 30 to 60 trees per hectare. That had risen to 2,000 in the 1990s. But higher concentrations of trees result in more intense fires that can consume even large

trees completely, rather than the more desirable ground fires that leave them intact.

At least five of the remaining 15 Woolsey plots were badly damaged this summer. Two in Arizona's Fort Valley Experimental Forest, the oldest US forest-service research station, sustained heavy damage. One portion was "completely consumed," says David Huffman, a researcher at Northern Arizona University, Flagstaff, who inspected the area after the fires were brought under control.

Plots near Los Alamos and Las Vegas, New Mexico, are also likely to have been at least partly destroyed, although scientists have yet to assess the most remote regions.

The damage hits hard because few plots have detailed records stretching back 100 years. "We've lost a long-term historical record," says Margaret Moore, professor of forestry at Northern Arizona University.

Fires have also damaged experimental plots in Florida and New Mexico that are part of a larger study of the best methods to restore forests to the conditions documented when the Woolsey plots were established.

The project involves surveying three to four blocks of land at 11 sites around the country. Each block will feature four treatments: control, manual thinning, controlled burning, and thinning and burning. The aim of the project is to find the best ecological approach to reducing the threat of large fires.

In a sense, the project has courted its losses, as the plots are placed in the kinds of forest most threatened by fires. "We are risking losing our experiments to the problems we are trying to measure," says Jim McIver, the project's director.

► <http://www.for.nau.edu/ecorest/woolsey.htm>

Electricity debt could mean lights out for Russian virus institute

Carl Levitin, Moscow & Alison Abbott

The plug will literally be pulled on the D. I. Ivanovsky Institute of Virology in Moscow next week if the institute cannot pay its 800,000 roubles (US\$30,000) debt to its electricity provider, Mosenergo.

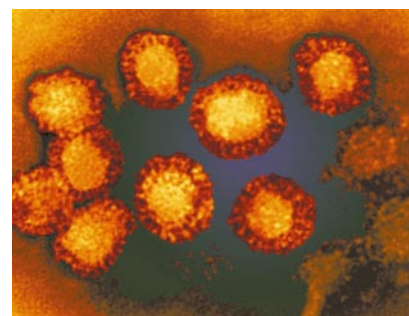
The institute's directors claim that much of its unique collection of viruses will be lost within hours of the electricity shut-down, as its supply of liquid nitrogen will not stretch to saving all strains. If this happens, another Russian scientific treasure will have fallen victim to the country's economic difficulties.

The institute's only income is from the Russian Academy of Medical Sciences. And, according to the institute's deputy director Petr Deryabin, it has no hope of raising the money by the 12 September deadline.

Mosenergo has already extended the original deadline by two weeks in acknowledgement of "the uniqueness of the institute". The debt has been accumulating since the beginning of last year.

Government departments have refused to help out, and the institute's international grants can only be used for designated research work, not for paying energy bills. "No one will be able to do any work without electricity, and the institute will have to close," says Deryabin. Moreover, he adds, it will be difficult to maintain safety conditions for containing the viruses.

The virus collection — more than 6,000 strains from 18 families — has been gathered over the past 60 years from around the former Soviet Union and through gifts and exchanges from virologists abroad. "It is an important collection historically and geographically," says Steve Elie, head of pathobiology at the chemical and



Doomed? The store of West Nile virus, shown here, would be lost without liquid nitrogen.

biological defence laboratories of Britain's Ministry of Defence.

Brian Mahy at the US Centers for Disease Control and Prevention says the threat is "extraordinary" considering the institute's value. The Ivanovsky Institute holds important collections of arthropod-borne viruses, he says, some of which have been collected from remote parts of the former Soviet Union and will be impossible to replace.

Particularly important, he says, is the institute's collection of the mosquito-borne West Nile virus. West Nile disease broke out in Russia a couple of years ago and reached the United States for the first time last year, causing five deaths in New York. The institute also has the world's best collection of tick-borne encephalitis viruses, says Mahy.

Despite facing the same difficulties as all Russian public research institutes, the institute has remained active. For example, it is a reference centre for monitoring influenza and other viral diseases for the World Health Organization.

The institute's director Dmitry Lvov, says that the collection "is one of Russia's real treasures and the country is going to lose it. It is a real tragedy." Lvov has asked various government departments for help. But he says the only response was from the Russian sanitary office, which sent a letter of protest to Mosenergo. ■

Researchers practise martian living in the Canadian Arctic

Steve Nadis

To prepare for life on Mars, a group of US researchers is exploring the next best thing — Devon Island, a barren, windswept outpost high in the Canadian Arctic.

Last month, investigators used a simulated martian habitat as a base for exploring the region. The aim is "to learn how to operate on Mars", explains Robert Zubrin, president of the Mars Society, a private organization funding the project.

So far, \$300,000 has been spent, and additional money has been raised to support operations up to the end of next year. A five-year, \$1 million research programme is planned to complement a NASA programme begun on Devon Island four years ago.

Things got off to a shaky start in July, when one of five payloads dropped onto the island by parachute crashed to the ground, destroying floor panels for the habitat, a crane and a trailer.

The team recovered by bringing plywood from the nearest town and using pulleys and winches instead of a crane. "When we get to Mars, we'll have to be prepared to improvise," says Zubrin. "Ultimately, human ingenuity rather than technology will carry us through." ■



The six-person habitat, a 7.2-metre-high cylinder with a domed roof, was finished in late July, two weeks behind schedule, allowing for a shortened occupancy period of less than a week. Next summer, researchers will spend two months at the base.

The crew, which included NASA scientists, simulated operations on Mars, spending 20 minutes in an airlock and wearing spacesuits. They used all-terrain vehicles and communicated daily with 'mission control' centres in Houston and Denver. ■

Californian centre will test medical uses of cannabis

Rex Dalton, San Diego

Marijuana is about to come under close scientific scrutiny in California. The Center for Medicinal Cannabis Research is being set up by the University of California to study the drug's efficacy and safety.

Announced last week, the centre will study the effects of cannabis on a number of diseases, including AIDS, muscle spasticity and nausea caused by cancer treatment. Some biochemistry might also be funded.

California's state assembly has provided US\$3 million for the first year of research, and plans to fund two more years at \$3 million each. Any research institution in California can apply for funds from the centre, which will be a venture between the university's San Diego and San Francisco campuses. Research proposals will be reviewed by a national panel of experts and approved by the National Institute on Drug Abuse, which will also supply the cannabis.

The plan is to start enrolling patients and conducting studies in January, according to Igor Grant, a physician at the University of



Munchies: a disabled man buys marijuana muffins at a cooperative in Oakland, California.

California at San Diego and co-director of the centre. He says that clinical trials and studies should provide reliable evidence on whether any medicinal benefits can be derived from smoking marijuana or ingesting marinol, a synthetic version of marijuana's most active ingredient, THC (delta-9-tetrahydrocannabinol).

Although there are surveys and

retrospective studies on the purported positive medicinal effects of cannabis, there are few prospective, blinded studies on the drug. US federal authorities have balked at funding such studies, concerned that they might encourage abuse of marijuana.

But in recent years, several prestigious scientific groups have called for a comprehensive cannabis study. Earlier this year, the National Academy of Sciences' National Research Council issued a report recommending research on cannabis.

California has voted to make cannabis legal for medicinal use, and 'cannabis clubs' provide marijuana for various purported health needs. This has prompted legal battles between state authorities and federal prosecutors, who fear that medicinal use will be used as a cover for criminal activity.

As these disputes went on, California state Senator John Vasconcellos, a Democrat from Santa Clara, successfully pushed the bill to fund the cannabis research centre through the state assembly. ■

♦ <http://ucsdnews.ucsd.edu>

Climate change expert stirs new controversy

Paul Smaglik, Washington

Opponents of the Kyoto Protocol on climate change have found fresh ammunition in the shape of a new scientific paper. The controversial article suggests that gases other than carbon dioxide are mainly to blame for the rapid global warming seen over the past few decades.

The Kyoto Protocol aims to reduce global warming by cutting CO₂ emissions, mainly from the use of fossil fuels. But the paper, from a team led by James Hansen of the Goddard Institute for Space Studies in New York, proposes that reducing emissions of methane, soot and the gases that cause photochemical smogs would be the easiest way to limit climate change in the short term.

Other climatologists have questioned the assumptions on which Hansen's conclusions are based, and note that the paper, published last week in *Proceedings of the National Academy of Sciences* (97, 9875–9880; 2000), has not been subjected to formal peer review.

Hansen helped alert the world to global warming in 1988, warning the US Congress: "It is time to stop waffling... the greenhouse effect is here." Not surprisingly, his latest work has been greeted with glee by the Global Climate Coalition, a Washington-based industry body that is lobbying against US ratification of the Kyoto Protocol. Its website concludes that "too much emphasis has been placed on the effects of burning fossil fuels".

Hansen's modelling assumes that levels of CO₂ in the atmosphere will continue to increase at a rate of 1.5 parts per million per year. He also calculates that tiny particles entering the atmosphere as a result of fossil-fuel burning — in particular aerosols of sulphates — have a cooling effect that is largely counterbalancing the warming caused by CO₂. Given this, Hansen argues that measures to limit global warming should focus on pollutants giving rise to ozone in the atmosphere.

But Hansen's assumptions are controversial. First, population growth and industrialization in the developing world mean that CO₂ emissions could grow. "I think it's actually very difficult to keep CO₂ growth rates to Hansen's targets," says Tom Wigley of the National Center for Atmospheric Research in Boulder, Colorado.

Second, huge uncertainty surrounds the size of the cooling effects of sulphates and other aerosols (see News Feature, pages 10–12). Most climatologists say that, in the absence of better data, we cannot assume that aerosols will offset the warming influence of CO₂.

Hansen argues that population growth and industrialization do not pose insurmountable problems. He says that renewable energy, and policies encouraging energy



Heating up: would reducing the methane emitted by rice fields make an impact on global warming?

efficiency, could help meet his predictions.

Michael MacCracken, director of the National Assessment Coordination Office of the US Global Change Research Program, agrees that targeting methane and soot could help stem global warming — at least in the short term. But he takes issue with Hansen's assumptions about fossil-fuel burning. Measures to clean up industrial emissions by removing sulphur would reduce cooling by aerosols, but leave the warming effect of CO₂ unaffected, MacCracken notes.

In stressing the importance of reducing methane emissions, for which rice cultivation is a major source, and of soot, Hansen's paper also puts fresh emphasis on the

need for action by developing countries. "Although the paper doesn't say so explicitly, it seems to shift the burden away from the West," says MacCracken.

But some US government advisers stress that limiting CO₂ — for which the United States is the biggest polluter — remains important. "Reminding everybody that particles and non-CO₂ greenhouse gases are important is fine," says John Holdren, a specialist on environmental policy at Harvard University, who chairs a panel that advises President Bill Clinton on energy issues. "But it's all too easy to get the impression from the article that CO₂ is not as important as has been thought, and that is not correct." ■

US grant glues 'virtual cell' together

Paul Smaglik, Washington

The US National Institutes of Health (NIH) this week launched a new series of grants for large-scale collaborative projects by pledging \$25 million over five years towards the construction of a 'virtual cell'.

The first ever 'glue grant' — so-called because it aims to link research between different institutions — has been awarded to the Alliance for Cellular Signaling (AFCS), a consortium based at the University of Texas Southwestern Medical Center and headed by Al Gilman, professor of pharmacology there.

The AFCS was formed last year to draw up a complete map of the interactions between one thousand proteins in two types of mouse cell (see *Nature* 402, 219; 1999). It links 50 investigators, who will provide planning and data analysis, seven laboratories devoted to conducting

experiments in protein–protein interaction, and 250 'members' who will specialize in compiling protein data and posting them on websites.

The alliance will initially screen for single protein–protein interactions, then build models of complete signalling networks once every element has been characterized. "We'd really like to represent these pathways completely," says Gilman.

The five-year project will actually cost \$50 million. Gilman is raising the other half from pharmaceutical companies and non-profit research organizations. The grant to the AFCS is one of several massive collaborative projects the NIH intends to launch during the next fiscal year that are intended to "facilitate the next evolutionary stage of integrative biomedical science". ■

♦ <http://afcs.swmed.edu>

CERN considers chasing up hints of Higgs boson

Alison Abbott, Geneva

Physicists at the European Laboratory for Particle Physics (CERN) in Geneva were locked in a dilemma earlier this week: whether to gamble SFr30 million (US\$17 million) following up results that could represent the first viewing of CERN's current holy grail, the elusive particle known as Higgs boson.

The data come from four different experiments done over the past six months using CERN's Large Electron-Positron (LEP) collider. But the LEP is scheduled to be switched off at the end of this month to allow construction of the Large Hadron Collider (LHC).

Keeping the LEP online would delay the LHC, possibly until 2006. This could give US scientists at Fermilab's Tevatron proton-antiproton collider time to pip CERN physicists at the post — particularly if the preliminary results provide a strong clue about the energy range to concentrate on.

Due to come online in 2005, the LHC will operate at energy levels of up to 14×10^{12} electron volts — high enough to detect the

Higgs boson with certainty, if it exists.

"A number of events have been observed in LEP experiments that are consistent with the production of the Higgs boson, in combination with the Z boson," says Roger Cashmore, CERN's director of research. But he points out: "Background processes throw up similar events, so more data will be needed to decide whether they are real."

The bulk of the data were collected at between 206×10^9 and 207×10^9 electron volts, the highest energy at which LEP operates stably. The results gathered so far must be compelling enough to persuade CERN's research board that two more months of experiments would "turn the indications into a recognizable real discovery that would be accepted by the scientific community".

"We don't want to be the scientists who go down in history as having missed Higgs," says Tiziano Camporesi, head of CERN's DELPHI experiment, who is convinced that there is a strong case for taking the gamble. A firm decision was due on Tuesday evening (see *Nature's* website for latest details) ■

► <http://www.nature.com>

news in brief

Treatment of accused physicist distresses US academics

Washington The presidents of the US National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine last week wrote to US Attorney General Janet Reno criticizing the government's treatment of Wen Ho Lee, the Taiwan-born physicist accused of stealing nuclear weapons secrets from Los Alamos National Laboratory.

Although Lee admits downloading information from Los Alamos computers, the presidents say that they are "distressed" over possible human rights violations. "We are concerned that inaccurate and detrimental testimony by government officials resulted in Dr Lee needlessly spending eight months in prison under harsh and questionable conditions of confinement," the presidents wrote.

UK universities must meet postgraduate standards

London A far-reaching review of UK university research says that universities will have to meet minimum standards for postgraduate

training if they are to receive funding. The report, by the Higher Education Funding Council for England, also says that some of the public money that it allocates to universities will go towards helping institutions meet strategic local, regional and national needs.

The report endorses the current system of giving institutions core support on the basis of excellence. But it says that the research assessment exercise used to determine this allocation should be revised to recognize "greater diversity in the characteristics of research excellence".

Lizard fossil returns to US museum

San Diego A fossil of a gliding lizard from the late Triassic is going from the auction room back to the American Museum of Natural History in New York. A California businessman, Dick Spight, bought *Icarosaurus siefkeri* at an auction in San Francisco last week for \$167,500, and promptly announced he would give it to the museum.

The 200-million-year-old specimen was found in a New Jersey quarry in 1960, given to the museum for identification, and reclaimed a decade ago by Alfred Siefker, one of its discoverers. In a move that upset many palaeontologists, Siefker recently put it up for auction, reportedly to pay medical bills.

Germans unearth hoard of Neanderthal remains

Munich An archaeological gold-mine has been uncovered at a site near Düsseldorf where the original Neanderthal man was discovered in 1856. German archaeologists say they have found three dozen fragments of bones in two caves near the Düssel river. Many are from the original Neanderthal man, including a piece of bone that had been missing from his skull.

They believe that the other bone fragments come from a second skeleton, around 44,000 years old, and suggest that this could represent one of the oldest examples of anatomically modern man. The find, announced last week, also includes thousands of stone tools and chips from tool production from Neanderthal



Missing links: archaeologists Jürgen Thissen (left) and Ralf Schmitz show the Neanderthal skull.

(50,000–40,000 years ago) and Cro-Magnon (around 30,000 years ago) times.

Also found at the site were burnt and cut bones of non-cave-dwelling animals. The archaeologists believe this is evidence of food preparation and cooking, indicating that the Neanderthals belonged to a settlement.

The original Neanderthal site had been located in a working quarry, but its coordinates were lost almost immediately after the skeleton was removed, and was only rediscovered in 1997.

Multidisciplined head appointed at CNRS

Paris Europe's largest basic research agency, France's Centre National de la Recherche Scientifique (CNRS), has a new director-general, Geneviève Berger. She replaces Catherine Bréchnignac, the first woman head of the CNRS, whose mandate ended in July.

Whereas Bréchnignac, like most past heads of the CNRS, was a physicist, Berger has taken an interdisciplinary path — she did a PhD in physics, and has worked in biology, nuclear medicine and imaging. Her laboratory produced the first commercial ultrasound bone imager, used to monitor osteoporosis. Berger comes from the science ministry, where she has been director of the technology department since January.

Trouble in the greenhouse

Tiny airborne particles affect the Earth's climate, in part by influencing the formation of clouds. But modelling the effects of these aerosols is proving to be one of the thorniest problems in climatology, says Mark Schrope.

"If I wake up with a nightmare, it is the indirect aerosol effect," says Veera-bhadran Ramanathan of the Scripps Institution of Oceanography in La Jolla, California. Ramanathan studies how the Earth's climate is influenced by atmospheric aerosols — dust, soot and other tiny particles. His research has helped reveal just how poorly their effects — in particular the 'indirect' effects mediated by cloud formation — are understood. As the world struggles to predict and control the extent of global warming, and to mitigate its effects, Ramanathan's nightmares could come to haunt us all.

Aerosols enter the atmosphere from a variety of natural and anthropogenic sources. Factories and diesel engines can pump out soot and sulphates; volcanoes and wildfires, such as those that have scorched the United States this summer, add to the haze. Meanwhile, winds whip up dust from deserts, and salt from sea spray. When they are carried into the atmosphere, these aerosols exert a direct cooling effect, reflecting sunlight back into space. But they also exert an indirect effect, influencing the formation and lifetime of clouds — and it is this that represents the biggest question mark in climate modelling. Most climatologists believe the overall influence of these indirect effects will be cooling, but that is about as far as they are prepared to go. "I don't think we have error bars yet," says Yoram Kaufman, an atmospheric scientist at NASA's Goddard Space Flight Center in Greenbelt, Maryland.

Removing these uncertainties will mean collecting extensive data to fill the vast gaps in current knowledge of the composition and global distribution of aerosols. To that end, a number of major satellite and ground-based projects are already underway or in the works. On the modelling side, all agree that a daunting task lies ahead, but the mood is one of cautious optimism. For the first time, researchers are expressing some confidence that they have identified the major components that must be included in models assessing the overall effects of aerosols. "I think we're at a crossroads," says Joyce Penner, a modeller at the University of Michigan in Ann Arbor. "Now the challenge is to figure out how to really parameterize these things correctly."

One of the main complicating factors in modelling the effect of aerosols is their short residence times in the atmosphere. Typically, particles remain aloft for a week or less, being either dragged down by their weight or removed by rain. In contrast, molecules of carbon dioxide persist for about a century,

and other greenhouse gases also have long residence times. So although greenhouse gases become well mixed on a global scale, levels of aerosols can vary widely on scales of less than one kilometre. And because concentrations at a given spot are highly influenced by winds, air masses loaded with aerosols can be carried a long way from the sources of the particles.

Up in the air

The chemistry of aerosols adds to the complexity. Many aerosol particles begin as gases. For example, sulphur dioxide coughed out of a volcano or smokestack can be oxidized to form particles of sulphate salts. But the formation of these sulphates is influenced by ozone — and so is affected by the presence of photochemical smogs. As a result, changes in atmospheric chemistry arising from pollution controls could influence the effects of aerosols on climate in ways that are difficult to predict. "It's a complicated system," says John Seinfeld, an atmospheric scientist at the California Institute of Technology in Pasadena.

Direct aerosol effects are relatively straightforward to calculate because the ways in which various particles scatter and reflect sunlight are well known. Even so, there are complicating factors — one of which was recently documented by the Indian Ocean Experiment (INDOEX), an international study of aerosols and atmospheric chemistry conducted in 1998 and 1999 that integrated ground-, ship- and aircraft-based measurements with satellite data. INDOEX revealed that dark particles such as soot can have a warming effect by absorbing solar energy^{1,2}. "We just made the problem a lot more complex," says Ramanathan, who was co-chief scientist for the project with Nobel laureate Paul Crutzen of the Max-Planck Institute for Chemistry in Mainz, Germany.

Calculations of direct aerosol effects can be checked by correlating data on aerosol concentrations with satellite measurements of reflected solar radiation. Model-derived estimates of direct effects of aerosols being considered by the Intergovernmental Panel on Climate Change (IPCC) are still in error by as much as 50% when compared with these data. But the models can predict general patterns of highs and lows in direct aerosol-induced cooling, says Penner, who is lead author of a chapter on aerosols for the next IPCC report, due to be approved at a meeting in Shanghai in

Indirect cooling effects of aerosols could be similar in size to the warming effects of greenhouse gases.



A question of balance: anthropogenic aerosols might be counteracting greenhouse warming.

January 2001. "I wouldn't say we're doing perfectly," she says, "but it's very encouraging."

Integrating indirect aerosol effects into global climate models is still in its infancy. The modellers did not begin including these effects until after the last IPCC report in 1995, and their results so far have led to a dramatic decline in the certainty with which the overall effects of aerosols can be predicted. "Indirect forcing is potentially a pretty important thing, and we really do need to learn how to model it," says Penner.

Blowing hot and cold

The indirect effects fall into two main categories. First, aerosols can act as cloud condensation nuclei, spurring the formation of more water droplets in a cloud than would normally be present. This increases the reflectance of clouds, causing more solar radiation to bounce back into space — and so should exert a cooling influence. This concept was first put forward by Sean Twomey of the University of Arizona in the 1970s³, but received little attention until 1987, when the effect was observed for aerosols in emissions from ships⁴. The Twomey effect has since been included in some climate models, but with little reliability.

The second category is rain suppression. When more droplets form within a given cloud, the available water is spread more sparsely, making each droplet smaller. This can prevent the droplets from growing large enough to fall as rain, increasing the cloud's lifetime and potentially prolonging its cooling effects. Bruce Albrecht of the University of Miami suggested in 1989 that aerosols could suppress rain⁵. Recently, Daniel Rosenfeld of the Hebrew University of Jerusalem showed this to be true for aerosols from burning vegetation⁶ and industrial pollution⁷, using data from the Tropical Rainfall Measuring Mission satellite. Researchers led by Maria Cristina Facchini of the Institute of



Unknown quantity: volcanoes, sea spray and wildfires all add aerosols to the atmosphere, but predicting their effects is proving to be difficult.

Atmospheric and Oceanic Science in Bologna, Italy, have also shown that organic aerosols dissolving in water droplets reduce the surface tension, which in clouds would inhibit droplet growth, preventing rain⁸.

But as with direct aerosol effects, INDOEX revealed an indirect effect that would cause warming: the absorption of heat by soot can 'burn' away cumulus clouds⁹.

Numerous other studies have added to the uncertainty. For instance, Qingyuan Han and co-workers at the University of Alabama in Huntsville have found that in certain types of clouds, decreasing droplet size decreases the clouds' reflectance¹⁰. Aircraft emissions of aerosols also complicate the picture when they trigger the formation of ice clouds high in the atmosphere. These clouds reflect less solar radiation than those containing water droplets and can also absorb heat, making their overall influence one of warming¹¹. Integrating these effects into current climate models will be difficult. The problem is that the models divide the atmosphere into 'cells' with a scale of around 100 kilometres, whereas modelling clouds would ideally use scales closer to one kilometre.

Although the uncertainties remain large, some modellers predict that indirect cooling effects could be similar in size to the warming effects of greenhouse gases^{12,13}. If so, that raises the troubling possibility that aerosols



could be counteracting the full effect of greenhouse warming. This could mean that the coming decades will see even more calamitous climate change than current models suggest — particularly if clean-air controls result in a reduction in anthropogenic aerosols. Whatever the truth, the uncertainty over aerosol effects makes it difficult to determine how, exactly, humans have altered the climate. “It means that we should be cautious in interpreting the past 100 years of climate change,” says Penner.

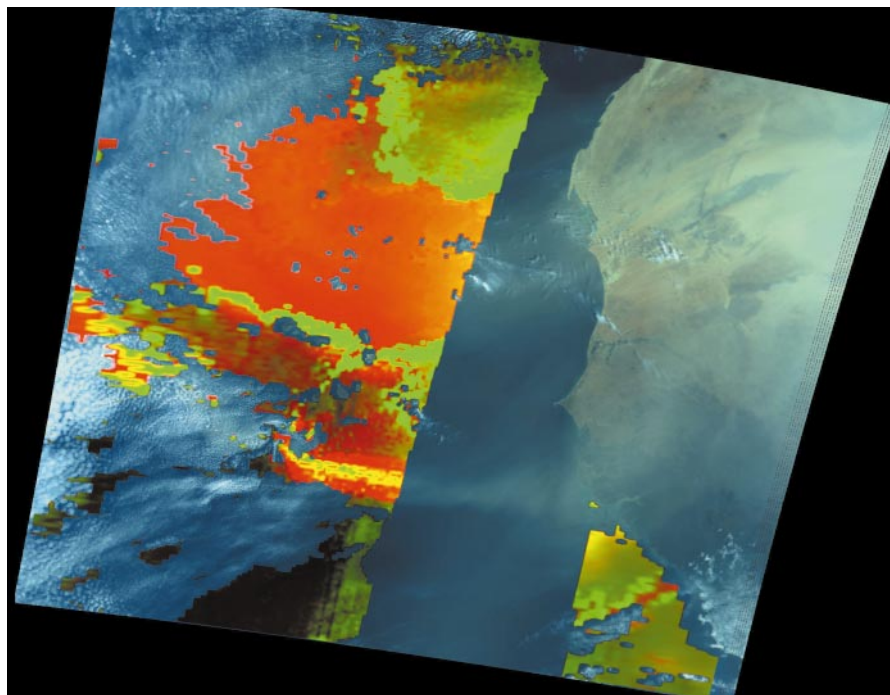
In the past, opponents of stringent action to combat global warming have seized on such scientific uncertainties. But most experts say that delaying measures to reduce greenhouse emissions until the indirect aerosol effect is better understood would be a grave mistake. “That debate should proceed independently of any consideration of particles,” says Seinfeld, “because those greenhouse gases are with us for a long time.”

Terra's firmer data

Nevertheless, answering the questions surrounding the indirect aerosol effect remains a top priority for climatologists. “I think we’ve got our work laid out for us, but even the next five years could see some much better answers,” says Penner. That prediction is based in large part on the anticipation of new data that should soon begin pouring in from satellites. The chief source will be the international satellite Terra, launched in December 1999, which is the flagship of NASA’s Earth Observing System.

Three of Terra’s five instruments will provide an unprecedented global view of aerosols. The Moderate-Resolution Imaging Spectroradiometer, or MODIS, measures reflected solar radiation at 36 different wavelengths across a 2,000 kilometre-wide swath, giving near-full global coverage once a day. This is providing the first global view of particle size distributions, because larger particles, such as mineral dust and salt, reflect more radiation at longer wavelengths. An instrument called MISR, the Multi-Angle Imaging Spectroradiometer, takes radiative measurements at four wavelengths and nine different angles giving a three-dimensional view of the atmosphere. Near major sources of aerosols, such as plumes of smoke, it will show the vertical distribution of particles in the atmosphere.

To round out the picture, CERES, or the Clouds and the Earth’s Radiant Energy System, will measure reflected sunlight and thermal radiation to determine the overall effect on climate of the aerosols present. Direct aerosol effects will be easier to study than indirect ones, as there is still no way to profile aerosols within clouds. But Kaufman, who is the project scientist for Terra, says that in some cases aerosols within clouds can be inferred from measurements of the clear air at the same altitude. He also hopes that it will prove possible to devise ways of using Terra’s



Homing in: an image of dust blowing from the Sahara to the Atlantic Ocean taken by NASA’s Terra satellite. Analysis of the data reveals mineral dust (red) and aerosol particles (green).

data to determine the proportion of atmospheric aerosols that are anthropogenic. “That’s something for the future,” he says.

Terra will be followed into orbit in December by a satellite called Aqua, which has similar capabilities but will fly in a ‘later’ orbit. This will, for example, give it an afternoon view of the tropics, compared with Terra’s morning flyover. This is important as factors such as cloud cover vary with the time of day. In 2003, NASA and CNES, the French space agency, intend to launch PICASSO-CENA (Pathfinder Instruments for Cloud and Aerosol Spaceborne Observations-Climatologie Etendue des Nuages et des Aerosols). Flying in formation with Aqua, this will use lidar — the laser equivalent of radar — to study the vertical distribution of aerosols and clouds.

Satellite measurements will be coordinated with those from a network of about 100 automated ground-stations known as the Aerosol Robotic Network, or AERONET, run by NASA, CNES and CNRS, France’s national research agency. AERONET will provide better data on the size distribution of particles — a key factor in modelling cloud condensation nuclei. Terra and the other satellites also perform best over water where background reflection from the Earth is relatively constant. AERONET will do a better job of measuring aerosols over land. It will also be used to validate satellite measurements.

The Asian-Pacific Regional Aerosol Characterization Experiment (ACE-Asia), which will combine satellite, aircraft and surface-based measurements in a similar manner to INDOEX, should also be in full swing by next year. Pollution from Asia is a major source of aerosols, and prevailing

winds carry much of them over the Pacific Ocean. ACE-Asia will provide the first full characterization of these aerosols. “They have the most complicated soup that you could imagine coming off Asia,” says Seinfeld, who is involved with the project.

When data from these projects are placed in the modellers’ hands, climatologists are optimistic that indirect aerosol effects will no longer seem such an intractable problem. “Some of the best minds in the business are working on this,” says Seinfeld. “It’s going to be a while, but I think that we’re going to understand it sooner or later.”

Mark Schrope is a freelance writer in Richmond, Virginia.

1. Satheesh, S. K. *et al.* *J. Geophys. Res. Atmos.* **104**, 27421–27427 (1999).
2. Satheesh, S. K. & Ramanathan, V. *Nature* **405**, 60–63 (2000).
3. Twomey, S. J. *J. Atmos. Sci.* **34**, 1149–1152 (1977).
4. Coakley, J. A., Bernstein, R. L. & Durkee, P. A. *Science* **237**, 1020–1022 (1987).
5. Albrecht, B. *Science* **245**, 1227–1230 (1989).
6. Rosenfeld, D. *Geophys. Res. Lett.* **26**, 3105–3108 (1999).
7. Rosenfeld, D. *Science* **287**, 1793–1796 (2000).
8. Facchini, M. C. *et al.* *Nature* **401**, 257–259 (1999).
9. Ackerman, A. S. *et al.* *Science* **288**, 1042–1047 (2000).
10. Han, Q., Rossow, W. B., Chou, J. & Welch, R. M. *J. Clim.* **11**, 1516–1528 (1998).
11. Intergovernmental Panel on Climate Change *Aviation and the Global Atmosphere* (Cambridge Univ. Press, 1999).
12. Lohmann, U., Feichter, J., Chuang, C. C. & Penner, J. E. *J. Geophys. Res. Atmos.* **104**, 9169–9198 (1999).
13. Rotsteyn, L. D. *J. Geophys. Res. Atmos.* **104**, 9369–9380 (1999).

Web links

INDOEX ♦ <http://www.indoex.ucsd.edu>
 IPCC ♦ <http://www.ipcc.ch>
 Terra ♦ <http://terra.nasa.gov>
 Aqua ♦ <http://eos-pm.gsfc.nasa.gov>
 PICASSO-CENA ♦ <http://www.picassocena.larc.nasa.gov/picasso.html>
 AERONET ♦ <http://aeronet.gsfc.nasa.gov:8080>
 ACE-Asia ♦ <http://saga.pmel.noaa.gov/aceasia>

Citation figures suggest that the UK brain drain is a genuine problem

Sir—Although the UK government's announcement of substantial new investment in the British science base¹ was reported in the national media as being largely about stopping the 'brain drain' of researchers to the United States, previous reports had suggested that there was no firm evidence of a brain drain². Using bibliometric data, we report here a statistically significant difference between the quality of scientists who trained in the United Kingdom but are now in the United States, and those who have stayed in the United Kingdom.

We studied a 10% stratified sample of people who obtained a doctorate in a science subject from a UK university in 1988 (ref. 3). This cohort has had sufficient time to settle into tenured or tenure-track positions, and their early publications have had time to accumulate citations.

We checked each individual researcher against the Science Citation Index (SCI) of the US Institute of Scientific Information for 1985 to 1989 and recorded the number of first-authored scientific papers published during this period, and the number of citations that each article has since received up to and including May 2000, for each individual. We used the SCI to check the publication records of individuals from 1998 to 2000, from which we ascertained whether researchers are still active and, if so, the country in which they are based.

We analysed the data by the current country of abode to test the quality of work published a decade ago between those who are now in the United States, those in the United Kingdom and those elsewhere. Many of those in the original sample had not published an article in 1998–2000, and were omitted from the analysis. There are, of course, many reasons for not publishing, including name changes, working in industrial laboratories with non-publishing policies, career breaks, radically changing one's research area, and leaving research completely. We do not know the nationality of the scientists whose careers we have tracked, but official statistics show that people living in the UK account for a consistent fraction of approximately two thirds of people awarded doctoral degrees from UK universities⁴.

We found that of the 770 scientists in our sample who were awarded their doctoral degrees in 1988, 252 are known to be still publishing. This is a 33% continuation rate, and represents a minimum rate.

Of the 252 scientists we tracked, 157 (62%) have most recently published from a UK address, 43 (17%) from the United

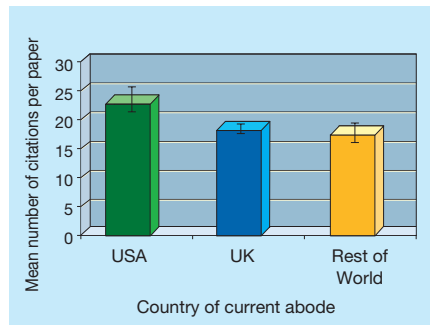


Figure 1 Mean number of citations per paper up to 2000, for papers published in 1985–89 by researchers with a doctoral degree from a UK university in 1988, analysed by country of current residence.

States and 52 (21%) from elsewhere.

The number of articles published from 1985 to 1989 is highest for scientists now publishing from UK addresses (2.40 ± 0.24 publications per person). This is not significantly higher than those now publishing from the United States (2.07 ± 0.43), but is significantly higher than those now publishing from the rest of the world (1.15 ± 0.24).

On the other hand, the mean number of citations per article for scientists who trained in the United Kingdom is significantly higher for those who are now in the United States than it is for those who remained. UK 'stayers' have more, but not significantly more, citations per paper than those currently publishing from addresses elsewhere in the world (see Fig. 1).

This is not, of course, proof of a brain drain. There are many other indications of quality than the number of citations a paper receives, and a scientist is more than the sum of his or her publications. A more thorough study of the same area is needed, which would include the nationality of individuals, a larger sample, a longer time-period, additional quality indicators and a reverse study. But our analysis so far does show that the UK government is right to regard the brain drain as a serious issue.

Alice Sharp Pierson, Peter Cotgreave

Save British Science Society, 29–30 Tavistock Square, London WC1H 9QU, UK

1. Loder, N. *Nature* **406**, 335 (2000).

2. *Financial Times* 1 July (2000).

3. *Index to Theses* (Aslib, London, 1990).

4. *SET Statistics 1999: A Handbook of Science, Engineering and Technology Indicators* (Department of Trade & Industry, London, 1999).

Debt and poverty turn a disease into an epidemic

Sir—Many of the scientists attending the Durban conference on AIDS^{1–4} have warned that South Africa's AIDS crisis can be attributed in part to President Thabo Mbeki's consultation with 'dissidents'. But

are the president's words and those of his colleagues really products of denial and confusion supported by pseudoscience? President Mbeki's central point is clear: alleviating poverty and social inequalities will play a pivotal role in conquering the African AIDS crisis.

Scientists at Durban asked a reasonable question: if treatments are necessary to conquer AIDS, why hasn't President Mbeki purchased drugs? Certainly, AIDS drug costs have received considerable attention, but the lack of public-health infrastructure in Africa remains largely ignored. Why is health infrastructure so poor in these countries? The answer, again, links AIDS with poverty. World Bank figures indicate that average healthcare spending on an African person is US\$14 a year, compared with \$2,673 for the average US citizen.

In the 1960s, when public health was considered a priority by emerging independent states, investments in social spending were rewarded with declines in infant mortality and a corresponding rise in life expectancy. But by the 1980s, many African countries had plunged into an economic crisis, firmly entrenching current trends of increasing poverty, hunger and HIV infection.

Oddly, the international response to these concerns was to slash healthcare programmes as part of austerity measures called 'structural adjustment programmes' (SAPs), imposed by the World Bank and the International Monetary Fund. As the benefits of SAP-led economic growth failed to trickle down to the poor, conditions for HIV infection among native populations worsened. One example of SAPs creating the conditions that favour the spread of HIV infection is the displacement of young women to cities where they resort to commercial sex work⁵.

As AIDS establishes a stranglehold in South Africa, President Mbeki's government is struggling to pay more than \$24 billion to the World Bank, and thus remains unable to begin carrying out the public-health initiatives it has planned.

Given the strong correlations between poverty and AIDS, one might question the priorities of those scientists who have chosen to battle with dissident colleagues rather than making the main thrust of their campaign the epidemiology of a disease worsened by sustained indebtedness.

Sanjay Basu, Kedar Mate,

Paul E. Farmer

Institute for Health and Social Justice, Department of Social Medicine, Harvard Medical School, 643 Huntington Avenue, Boston, Massachusetts 02115, USA

1. Cherry, M. *Nature* **406**, 3 (2000).

2. *Nature* **406**, 15–16 (2000).

3. Cherry, M. *Nature* **406**, 113 (2000).

4. Cherry, M. *Nature* **406**, 223 (2000).

5. Lurie, P., Hintzen, P. & Lowe, R. *AIDS* **9**, 539–546 (1995).

Expert witnesses take the stand

Historians of science can play an important role in US public health litigation.

Robert N. Proctor

Science and the law have never been easy bedfellows. Both rely on evidence and reasoned judgement to arrive at truth, but the methods used and the end results are generally very different. For example, legal proceedings must arrive at a kind of satisfactory closure, but a scientific conclusion is at least theoretically provisional, and open forever to revision.

Experts have been called into court as witnesses for centuries: in France, for example, following Roman law, some of the earliest experts were handwriting analysts, brought in to distinguish genuine from forged signatures. The use of historians in the court, however, is rather new. Science historians are increasingly being asked to serve as expert witnesses, especially in cases of product litigation, where the question is often "who knew what when?" about the nature or extent of a hazard.

Contemporary knowledge

The reason science historians are in demand is that science does not invariably progress in a strictly linear or cumulative fashion. Scientific knowledge is accepted (or ignored) in different ways at different times and places, making it important to establish what was known concretely by specific individuals at a given place and time. It could be of great legal import to know, for example, whether the tobacco aetiology of lung cancer was first recognized in the 1930s or in the 1960s. Such a question is often at the heart of class-action suits seeking compensation for negligence.

Historians of science are particularly adept at making such assessments, especially when the question is how a theory or conviction moved from one community to another, filtered into public consciousness, and was accepted or rejected. In the courtroom, one often needs to provide testimony on when a scientific consensus was achieved, or whether one could have expected the public or a company to be aware of a particular hazard. This could cover such issues as whether a particular hazard (say asbestos or radiation emitted from mobile phones) was widely recognized or sharply contested; whether estimates of the magnitude of that hazard were based on mainstream or marginal ideas; whether steps taken to advertise the hazard were consistent with prevailing practices, and so forth.

I have been employed in three such cases during the past few years, the questions in each being: was it reasonable to have known at such and such a time that a particular substance or procedure was hazardous? And did the people responsible for causing the injury

in question act responsibly, given the scientific and ethical standards of the time?

The first was a fairly straightforward class-action suit brought on behalf of 2,800 women from Alberta, Canada, who were forcibly sterilized as part of that province's eugenics legislation, from the late 1920s until the early 1970s. I had originally been asked whether I could evaluate the science behind the legislation; I told the lawyers involved that this was probably not such a good line of comparison, since many other countries had had similar laws, and Alberta in this respect did not seem out of step with other parts of the world. I suggested instead that I might compare the Canadian legislation with the sterilization legislation of Nazi Germany — as Alberta had actually sterilized a higher proportion of its population than the Nazis, which surely would help the defence. I prepared a written testimony for the plaintiffs, who eventually won the case.

The second was the 'Vanderbilt case', a widely publicized class-action suit on behalf of 830 pregnant women who were fed radioactive iron from 1945 to 1947 as part of a 'nutrition experiment' sponsored by the Rockefeller Founda-

tion, the Atomic Energy Commission and Vanderbilt University. The key question here was whether the organizers and administrators of this experiment acted responsibly and in accordance with contemporary ethical standards. The university had already admitted that at least four of the women had given birth to children who had died at an early age from blood cancers, a rate judged statistically significant even by Vanderbilt university scholars (the expected value for this population was 0.6). None of the women had ever been informed that their children might have died as a result of their participation in the experiment.

The Vanderbilt case was complicated and involved many other kinds of experts than historians. My job was to explore whether the experimenters could have known about the potential hazards before administering the radionuclides. Dosage was one issue of dispute, as was the question of whether the 'zero tolerance' theory of dose-response was accepted, as one key question was whether the doses involved could reasonably have been seen as unsafe. I showed, for example, that the 'single-hit' theory of carcinogenesis, devel-



Emma Craft, whose daughter died of cancer aged 11, won compensation for the Vanderbilt experiment.

oped by Nikolai Timofeeff-Ressovsky and Max Delbrück in Germany in 1935, should have been familiar to the US investigators, as Delbrück had emigrated to the United States in the late 1930s and had ended up as head of the very department where Paul Hahn, the principal investigator of the Vanderbilt experiment, was employed. The plaintiffs won that case — a \$10 million settlement plus a formal apology from the president of the university to the women who had lost their children.

The third trial was a 1999 class-action suit on behalf of the pension funds of trade-union workers in Ohio, seeking damages for injuries to the membership caused by tobacco. One of the questions I was asked to address was how early the medical case against tobacco had developed and how the tobacco industry had responded. I also provided testimony on how quickly this was accepted within the US medical establishment, and how quickly the public had accepted these findings. A further question was whether the industry itself had been aware of the cancer hazard, and whether it had taken steps to inform consumers.

Important in this was the timing of the discovery and popularization of the tobacco hazard. I knew that German physicians had established clinical experimental evidence of a lung-cancer hazard in the 1930s and 1940s, that British and US scientists had nailed down the link in the 1950s, and that by the middle of the 1950s it was possible to talk about a 'scientific consensus' over the major dangers of cigarettes — especially lung cancer and heart disease.

I could also document, though, that the industry had devoted considerable time and money to convincing the public that the case against cigarettes was "merely statistical", that action against cigarettes should wait until "further research" proved the precise nature of the hazard beyond any shadow of doubt, and so forth. The tobacco industry spent hundreds of millions of dollars on purported 'research', one function of which was to insinuate doubt concerning the scientific consensus.

Though the tobacco industry for many years maintained that cigarettes were safe, its legal stance for the past few years has been that by the 1950s an informed consumer should already have been fully aware of the nature of the risk. So it was important for the plaintiffs to show that the industry played a crucial part in maintaining the idea that the scientific case against tobacco was still open. A central question was the role played by the industry in fostering public doubts about tobacco's hazards; the plaintiff's position was that, despite the scientific consensus, it took quite some time for that consensus to be adopted by the public, because the industry was doing everything it could to convince the public otherwise.

A further issue was whether any cancers and heart disease had actually been caused by the industry's failure to notify the public of what it knew about the nature of the risk. In my written testimony, and under cross-examina-

But how does one determine a historian's rate of error?

tion in the courtroom, I produced a quantitative estimate of how many cigarettes would not have been smoked in the United States had the industry admitted what it knew in a timely fashion. This was obviously a somewhat speculative activity, but not without historical foundation: I calculated that roughly 8 trillion fewer cigarettes would have been smoked from 1954 to 1994, based on the assumption that the decline in US per capita consumption which began in 1964 could and should have begun ten years earlier than it did. Hence it was possible to estimate how many lung cancers could have been prevented, given that one lung cancer will be produced for every 2 million to 4 million cigarettes smoked in a given society.

I was apparently the first historian ever to testify for the plaintiffs in a federal tobacco trial; the tobacco industry, by contrast, has often used historians in its defence, especially to show that knowledge of tobacco hazards was widespread in the 1950s and 1960s.

Status of expert evidence

The theoretical and practical questions raised by giving evidence as a historical expert are myriad and complex. There is also the question of bias, because the expert is paid. The sums can be huge: I got \$150 per hour for my work, but I know that one of the statisticians testifying for the defence in my tobacco case (a Harvard professor) was paid ten times this rate. Is the neutrality of experts compromised by the fact that they are highly paid? How does the expertise produced in a US court differ from that of European courts, where experts are hired by the judge and paid very little?

There is also the interesting question of what is or is not admissible. In Anglo-American law, witnesses are normally expected to state only facts, not to express opinions. But expert witnesses can express opinions, if those opinions are considered to have emerged from their expertise. Ordinary witnesses are confined to what they themselves have seen or heard; experts can draw from a larger fund of established knowledge, the logic being that their knowledge or skills can help determine matters of fact or perspective that cannot be obtained through direct observation. (They cannot, however, expressly cite the work of other scholars, as that would be 'hearsay'.)

In the United States, official rules of evidence require expert opinions to be consistent with established methods. Standards of admissibility for expert testimony have changed recently: the Frye rule of 1923 required only that the opinions expressed be consistent with established views of the rele-

vant expert community; this was modified in 1993 by the Daubert rule, which now allows judges to decide whether the methods used by experts are accepted within their relevant communities, considering whether they are peer-reviewed, have demonstrable rates of error, and so forth. Opinions differ on whether the new standard makes it harder to admit expert evidence. Observers seem to agree, though, that it gives judges greater power to rule out certain kinds of expertise (for example 'junk science') as inadmissible.

In the tobacco trial, one central question became whether the historical reconstruction I had offered was consistent with the kinds of history that historians actually write. Was it really standard practice for a historian to say how history might have been different, if circumstances had been different?

I argued that it was legitimate for historians to say how particular actions or policies had affected history; that implicit in any such attribution of causality was the assumption that human history is not predetermined in any deep sense; that human affairs could have turned out differently if the forces pushing us one way or another had been different. It was an older, outmoded, Marxist-positivist notion from the nineteenth century, I argued, that held that people are just puppets of their environments. The plaintiffs lost that case, though the lawyers who hired me say the case had probably already been lost in jury selection.

Role of historians

Historians may well come to play a growing role as expert witnesses in the US courts, especially in product liability, where it is often crucial to establish whether or not it was reasonable to have known that a product could have posed a hazard. Scientific practices change over time, as do perceptions of what is an acceptable risk; historians are often in a good position to judge how and perhaps even why such changes have occurred.

But there may come a time when courts will have to decide what kinds of historical expertise are admissible. The criteria developed for scientific expertise in the Daubert rule do not clearly apply to historians. Peer review might be used as a yardstick, but how does one determine a historian's rate of error? Historical methodologies are diverse, and it may be difficult to find agreement on what methods are to be regarded as conventional.

There is much debate today about the nature and limits of expert witnessing — perhaps an unavoidably politicized issue in the adversarial legal system. The more that courts recognize the value of a historical perspective, the more likely we are to see further struggles over the question of what should count as legitimate expertise. ■

Robert N. Proctor is in the Department of History, Pennsylvania State University, University Park, Pennsylvania 16802, USA.
e-mail: rnp5@psu.edu

Varieties of silly experience

Or how to keep your head when all around are losing theirs.

Voodoo Science: The Road from Foolishness to Fraud

by Robert L. Park

Oxford University Press: 2000. 240 pp.

£18.99, \$25

N. David Mermin

Since the mid-1980s Bob Park, professor of physics at the University of Maryland and director of the Washington office of the American Physical Society, has produced a weekly newsletter, *What's New* (14 years' worth of which are available online at <http://www.aps.org/WN/>). His primary focus is on the surrealistic drama endlessly taking place at the interface between government and science. Park's acid views on the insight and judgement that some of America's distinguished leaders bring to bear on the uses and support of science bring his audience of physicists both delight (for his wit) and despair (over the occasions for that wit). Nor is his interest in the public misunderstanding of science limited to the corridors of power. Wacky encounters with science by amateurs, entrepreneurs, journalists, and even apparently bona fide practitioners, are all fair game. The delight he induces in his sponsors is apparently not entirely unalloyed, for since March 1994, his remarks have concluded with this carefully negotiated disclaimer: "Note: Opinions are the author's and are not necessarily shared by the APS, but they should be."

Voodoo Science is a loosely interwoven, fascinating compendium of tales of human folly, presented in the same engaging style readers of *What's New* have come to enjoy and admire. Although I found the hopping back and forth between examples somewhat disconcerting, page by page the book is a source of great pleasure. Park is a superb storyteller, he writes gracefully with energy and charm, and the varieties of silliness he chronicles are fascinating. Most scientists will enjoy the book immensely.

But scientists are not the audience Park wants to reach. "I will, of course, be delighted if scientists read my book and find it entertaining, but it wasn't written for them." His hope is to educate those who "choose scientific beliefs the same way they choose to be Methodists, or Democrats, or Chicago Cubs fans ... What can we tell [such] people that will help them to judge which claims are science and which are voodoo?"

Park offers such readers an intermingling of cautionary tales and pronouncements on the nature of scientific knowledge. The tales, which are too anecdotal and judgemental to be called case studies, leave



Science or folly? Clockwise from top left, the criticized self-support experiment Biosphere 2, plans for the unbuilt New Age theme park Maharishi Veda Land, part of the International Space Station before launch in Russia, and New Age powerbead bracelets promising a wide variety of health effects.

few stones unturned, and although some fairly unsavoury specimens come scuttling out, some scientists may also be disconcerted to find him kicking rather close to home. Among his targets are perpetual-motion machines, homeopathy, cold fusion (which surfaces throughout the book, somewhat misleadingly, even in a chapter on perpetual motion), denials that human activity is affecting the Earth's climate, the International Space Station ("a make-work welfare program for the aerospace industry"), alleged benefits of wearing lots of magnets, extrasensory perception, the Office of Alternative Medicine at the US National Institutes of Health, precognition, scientists who celebrate the strangeness of science, power lines said to cause cancer, the reduction of crime by mass transcendental meditation, Ronald Reagan's Strategic Defense Initiative, gravity shields, official secrecy, touch therapy, Biosphere II, unidentified flying objects, vitamin O, and invocations of quantum mechanics to account for just about any New Age notion you like.

Discoursing more generally on the nature of scientific knowledge, Park divides its misuses, which he collectively terms "voodoo science", into pathological science (in which hopes of fame or fortune lead you into fooling yourself about what you have actually accomplished), fraudulent science (in which you become aware of your self-deception but try to cover your tracks), junk science (in which you devise arguments that, although not necessarily incorrect, are

designed to deceive or befuddle others, such as jurists or legislators), pseudoscience (which is grounded in superstition or political or religious conviction), and — later in the book — unimportant science (whatever has been proposed to date as a justification for building the Space Station).

He also offers an offbeat theory of good science as an "antidote" to a genetically determined mechanism for generating belief that would constantly drive us down wrong paths if left unchecked. Skating perilously close to junk science himself, Park maintains that superstition originates when "the chemical messengers of emotion cause the thalamus to bypass the sensory cortex and route the information directly to the amygdala".

Less idiosyncratically, he also explains the role of the placebo effect in alternative medicine, the mischief-generating power of a generalized form of Pascal's wager (your pet idea may be grossly improbable but if it's right the pay-off would be so great that you must press on), and the Texas sharpshooter fallacy (shooting at the barn and then drawing a target around the hole).

I suspect Park will be most effective with those of his intended readers who can't tell science from religion, politics or baseball, with his many stories of human folly and preposterous claims. These are particularly compelling when he allows the absurdities to speak for themselves with a minimum of interpretive intervention. But when he starts invoking laws of physics I wonder why he

book reviews

expects them to carry more weight with such readers than appeals to the Ten Commandments, Magna Carta or Three Strikes and You're Out.

I found the book weakest when Park discourses on the nature of good science. Although it is easy enough to grasp the absurdity of most of his voodoo examples, drawing up a set of general principles cleanly dividing science from its voodoo manifestations is not as straightforward a business as Park implicitly suggests. How to assess the validity of scientific claims in courts of law, for example, is a notoriously vexing issue. Park cites approvingly the idea of "allowing evidence based on scientifically valid principles to be admitted even if it does not represent a scientific consensus".

Although acknowledging that this gives "judges little guidance in how to decide whether or not evidence is [so] based", Park has nothing illuminating to say about the subtle and important question of why "scientifically valid principles" can often lead to the kinds of diametrically opposed positions that do contend in the absence of a "scientific consensus".

Like Irving Langmuir's famous 1953 lecture on "pathological science", *Voodoo Science* is a work that most scientists will applaud because of its wonderfully entertaining histories of deception and self-deception, but that many historians, philosophers or sociologists of science will dismiss as oversimplified and naive in its general pronouncements on the nature of science. In spite of this absence of consensus, I would say that both judgements are correct. ■

N. David Mermin is in the Department of Physics, Cornell University, Clark Hall, Ithaca, New York 14853-2501, USA.

Blowing away the smokescreen

The Science of Marijuana

by Leslie L. Iversen

Oxford University Press: 2000. 283 pp.

£18.99, \$29.95

Raphael Mechoulam

The use of *Cannabis sativa* stretches back to the distant past. The Assyrians, ancient Egyptians, Indians and Chinese used it in medicine, and perhaps at times for recreation and in religious rites. Its use continued for millennia and parallels that of opium. Yet modern research on cannabis has always lagged behind that on *Papaver somniferum*, the opium poppy. Whereas morphine was isolated from opium early in the nineteenth century, delta-9-tetrahydrocannabinol, the



Joint action: vilified as a recreational drug, marijuana can help sufferers of medical conditions such as multiple sclerosis.

active constituent of cannabis, was not fully identified until 1964. The first opiate receptors and endogenous opiates were described in the 1970s, but those of the cannabinoids were not discovered for another 20 years.

Today, however, cannabinoid research is a central topic of investigations in many fields, particularly in neuroscience. Two receptors have been identified: CB1, found mostly in the central nervous system, and CB2, present mostly in the immune organs. Two endocannabinoids — anandamide and 2-arachidonyl glycerol — have been isolated; antagonists and knockout mice for CB1 and CB2 receptors are available.

The endocannabinoids are involved in neuromodulation, in the immune system and even in reproduction. Cannabinoid receptors interact with the opiate, GABA, dopamine and glutamate systems. We have learnt much about the effects of endocannabinoids on short-term memory, pain regulation, neuroprotection, appetite and blood-pressure regulation. Clinical work has shown that the endocannabinoid system may be involved, either as a

(neuro)protector or as the villain, in diverse medical conditions such as Parkinson's disease, schizophrenia, Tourette's syndrome and even miscarriage. Anandamide, the more thoroughly investigated endocannabinoid, has been billed as heading "for the big [neurotransmitter] league".

These advances are admirably described and critically discussed by Leslie Iversen, an eminent pharmacologist with wide experience both in academia and in industry. Having also served on numerous official bodies dealing with drugs, he is in a position to deal objectively with the political and social as well as the scientific aspects of the cannabis issue. He presents the history, biology and medical aspects of cannabis and cannabinoids with a critical eye, and his style is superb.

Over the past two decades a strong grassroots movement to legalize cannabis as a therapeutic drug has led to political pressure in several countries; the measure has been approved in several US state referenda. Campaigners point out that it reduces vomiting after cancer chemotherapy, enhances the appetite of people with AIDS and is active against pain, glaucoma, multiple sclerosis and many other conditions. Numerous prestigious groups — among them a House of Lords committee — have looked at the scientific data. Iversen's own analysis of the evidence for its clinical effectiveness indicates that it is woefully inadequate by

modern standards: well-controlled clinical trials are urgently needed, he says.

His analysis of the 'medical marijuana' issues and of the trend towards legalization (or decriminalization) is even-handed. He supports the view of the editor of the *New England Journal of Medicine* that "a policy that prohibits physicians from alleviating suffering by prescribing marijuana for seriously ill patients is misguided, heavy-handed and inhumane", although he expresses his view in milder terms — he would allow compassionate use in some well-established conditions. In the United Kingdom, for example, he is in favour of changing its legal status to that

of a Schedule II drug, permitting doctors to use it on a named-patient basis. And he believes that if well-controlled trials yielded positive data "it would be hard for any

government to resist an application for the approval of cannabis or a cannabinoid as a medicine". Is he somewhat naive about politicians?

On the use of cannabis as a recreational drug Iversen is much more cautious. Its short-term adverse effects are mild — mostly changes in time and distance estimation. However, as was the case with tobacco years ago, we may not have yet seen the results of long-term use. Cannabis smoke, like that of tobacco, is of course carcinogenic. Does it represent a medical or social danger? The jury is still out, but the data are not encouraging. Recent work indicates that delta-9-tetrahydrocannabinol promotes tumour growth by inhibiting certain anti-tumour immunity. Yet pot-smokers do not seem to have excessive cancer rates. As Iversen points out, epidemiological surveys are clearly needed to follow future developments.

Iversen has written an excellent, well-balanced, non-specialist book covering both the biological and social aspects of cannabis. I strongly recommend it.

Raphael Mechoulam is in the Faculty of Medicine, The Hebrew University, Ein Kerem Campus, Jerusalem 91120, Israel.

Lights in the night sky

Great Comets

by Robert Burnham

Cambridge University Press: 2000. 228 pp.

£14.95, \$21.95 (pbk)

Jay M. Pasachoff

To be a 'great comet', according to the definition credited to David Hughes of the University of Sheffield, a comet must have a large nucleus and coma, a large active surface area, reach perihelion near the Sun, pass close to Earth, and provide good viewing opportunities for terrestrial observers. Halley's Comet meets the last of those criteria every 75 or 76 years, although its 1986 passage through the inner Solar System was not one of the best of its 30 known apparitions. Great comets in the past 50 years have included Arend-Roland of 1956, Ikeya-Seki of 1965, West of 1976 and Austin of 1990. Robert Burnham takes advantage of the passage of two such comets a few years ago, known informally as Hyakutake and Hale-Bopp, to present a *mélange* of description, photographs and analysis to a general audience.

From Burnham's distinguished amateur-astronomy career, and his stint as editor-in-chief of the popular glossy magazine *Astronomy*, one could predict that his book would

be interestingly and clearly written, well illustrated, and have some non-technical but wide-ranging descriptions of scientific results. These predictions are well met in the published book, which uses heavy stock to print about 100 photographs and other illustrations, most in colour.

The story of comet observation from early times is briefly told, only touching on the work of Tycho Brahe, Johannes Kepler, Isaac Newton and Edmond Halley. Galileo's first telescopic observations of a comet, made in 1618, led that great scientist to conclude, wrongly, that comets appeared in the Earth's atmosphere, like aurorae or Sun-dogs.

Friedrich Bessel in 1836 correctly suggested that the 'nongravitational forces' that affect the orbits and periods of comets were caused by rocket effects of jets of gas that were ejected from the coma.

The stories behind the two great comets of the 1990s are told individually. Burnham relates how Yuji Hyakutake found a first comet in December 1995, and then a second one, which was to become famous, about a month later. He provides a well-drawn diagram of the comet's orbit and a series of photographs that includes not only its interesting tail but also an image taken by SOHO (the Solar and Heliospheric Observatory) when the comet was too close to the Sun to be seen from Earth. It ejected great amounts of water molecules and dust; and because of its proximity to Earth-based radio telescopes, many molecules — including methanol, methyl cyanide, formaldehyde and hydrogen sulphide — were detected. Burnham also includes the satellite image that surprisingly detected X-rays coming from Hyakutake. However, the text discussion of this doesn't appear for another dozen pages, the victim of a general problem in correlating text and figures. Indeed, it is not for another 50 pages, in the Hale-Bopp discussion, that we read about the two major competing models to explain the X-rays — that they come either from collisions between heavy ions in the solar wind and particles in the comet's coma or from dust particles reflecting solar X-rays.

The Hale-Bopp story is related in friendly fashion, including the circumstances in which Alan Hale and, independently, Thomas Bopp saw a fuzzy region in Sagittarius, checked that it wasn't on a sky atlas and that it visibly moved, and reported the discovery. This passage through the inner Solar System shortened the comet's orbit, so that it will return in only 2,404 years instead of the 4,265 years of its previous elliptical path around the Sun. Hale-Bopp was notable not only for its overall brightness but also for the easily viewed distinction between its whitish, curving dust tail and its bluish, structured plasma tail. Those who know the comet only as a historical

book reviews

name will be able to follow the sequence in the photos. These include many with scenic foreground views boasting of the comet in the background sky. Exposure information is a useful guide to photographing new comets.

Burnham includes a chapter that not only surveys great comets of the past but also looks at the several space missions that met Hale-Bopp. It is not, incidentally, properly called "the Hale-Bopp Comet", the name that was used in the popular press and radio discussions preceding the suicides of the Heaven's Gate cult. Burnham relates

the story of the mistaken, hasty identification of a supposedly trailing object by an amateur astronomer and its premature description to a radio talk-show host. Although the object was soon identified as a normal, though faint, star that was not in the sky-mapping program used by the amateur astronomer, this explanation was not believed by cultists. I was interested to see reproduced the colour photograph showing the home page of the Heaven's Gate website.

The book ends with a discussion of near-Earth objects, from the Tunguska impact of

1908 to the Spaceguard plan to set up a network of automated telescopes to detect Earth-threatening asteroids. I hadn't known that spaceborne detectors reveal that, about ten times a year, there are kiloton-energy impacts of meteoroid material.

Burnham's book provides an overview of comet observation accessible to every reader, with enough indications of the science involved to show the vitality of the research. I recommend it to a wide range of readers. ■

Jay M. Pasachoff is in the Department of Astronomy, Williams College, 33 Lab Campus Drive, Williamstown, Massachusetts 01267, USA.

Science in culture

Mannerists' monsters

The More Than Meets the Eye exhibition at the Victoria and Albert Museum in London

Philip Ball

One of the objections to the idea that art and science share common principles is that science is cumulative whereas art is sequential. In science, new knowledge builds on and refines the old. In art, according to common caricature, the new eclipses the old and rewrites all the rules. Although it would be difficult in practice to find examples of art that were not informed by the past, one could nevertheless make a good case that artists have sometimes dealt with new technical knowledge in an arbitrary manner, accepting or discarding it as current fashion dictates.

Skills in handling paints as chemical substances gradually deteriorated between the Baroque era and the nineteenth century, and had catastrophic consequences for the longevity of some works. In a quite different context, the much-praised determination of Renaissance artists to paint "true to nature", which required a study of perspective, anatomy and proportion, did not even outlive some of its greatest masters. By the late sixteenth century, mannerist artists such as Parmigianino and El Greco were introducing bizarre and often grotesque distortions to the human form.

The relationship between mutation in nature and that introduced by artists is one of the themes explored in the *More than Meets the Eye* exhibition running during 6–30 September at the Victoria and Albert Museum in South Kensington, London. The exhibition, a collaboration between the Royal Society, *Nature* and the Victoria and Albert Museum, and sponsored by L'Oréal, is part of the "Creating Sparks" festival of science and art organized by the British Association. It takes the form of a self-guided trail through the museum, where ten displays explore science-related aspects of the collection.

The theme of mutations is suggested in part by the several chimaeras of human and animal in the collection of Indian sculpture. In Indian and Middle Eastern mythology such figures



Distorted view: the mannerists exaggerated the human form for dramatic effect.

commonly depict deities. In Greek mythology these chimaeras symbolize animalistic passions, more bacchanalian or lascivious than transcendental: the satyr, the minotaur, the mermaid. Either way, such hybridization is a badge of the supernatural, a proclamation that mixtures that cross the species barrier lie outside the natural order.

But the elegant 'monsters' created by the mannerists raise perhaps more pertinent questions about the extent to which science has informed art. Mannerism arose from a Renaissance debate about the nature of 'truth' and 'beauty' in art.

There was by no means universal agreement in Renaissance Italy about how the natural world should be portrayed. Most artists agreed with the Genoese architect and theorist Leon Battista Alberti (1404–72) that the foremost aim was to achieve beauty. According to Alberti, rational principles such as perspective were simply a means to this end.

Yet what is beautiful? To Alberti the rational humanist, it was almost a calculable quantity: "a kind of harmony and concord of all the parts to form a whole which is constructed according to a fixed number, and a certain relation and order, as symmetry... demands." Yet as nature was often imperfect, Alberti proposed that the artist select only its most pleasing aspects.

Leonardo da Vinci had a similar conviction that artistic judgement should be rational. But he saw little need to censor nature's 'ugliness', as his sketches of people with deformities show. He felt that the imperfect provided a foil to make beauty shine more brightly.

Michelangelo, on the other hand, pre-empted the decline of humanist objectivity into the arbitrary conventions of mannerism. Scientific principles, he maintained, were of little account in creating beauty: "all the reasonings of geometry and arithmetic, and all the proofs of perspective were of no use to a man without the eye." It is then but a small step to the elitist criterion of 'taste' (*grazia*) championed by mannerists such as Federico Zuccaro, president of the Academy of Drawing in Rome. In 1593 he pronounced that "the art of painting does not derive from the mathematical sciences, nor has it any need to resort to them to learn rules or means for its own art".

Mannerism became a period of wild experiment, of flashiness parading as originality to catch the courtier's eye. And science took a back seat.

Philip Ball is a consultant editor at *Nature*. ■

Old wine in new bottles

An agenda for biology that's in its second century — and still going strong.

Jane Maienschein

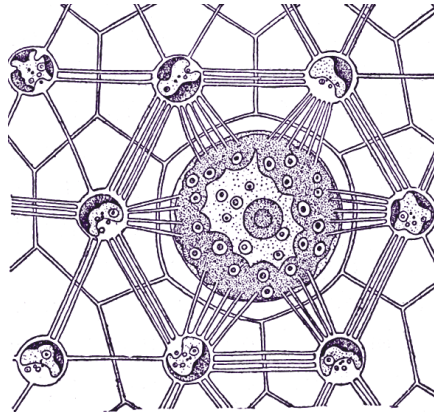
"The magnitude of the problem of development, whether ontogenetic or phylogenetic, has been underestimated," wrote Edmund Beecher Wilson. "Yet the splendid achievements of cell-research in the past twenty years showed us the promise of the possibilities for the future."

Wilson rejected both the theorists' claims to have answered the important questions about life and the vitalists' insistence that we can never answer them with materialistic science. He insisted that through experimental research, "step by step the way may yet be opened to an understanding of inheritance and development". The quotations are from *The Cell in Development and Inheritance*, published in 1896. Wilson's emphasis on a balanced study of all aspects of living organisms remains as compelling today, and reminds us of the value of studying the history of biology as well as of organic life.

In three books published during 1895–96, Wilson addressed what today we call biocomplexity. His positivistic programme for scientific progress was grounded in epistemology, epigenesis, evolution and education.

His epistemological values reflected the increasingly experimental approach to biology at the end of the nineteenth century. Wilson felt that science should seek "positive results" through detailed empiricism, rejecting metaphysics. The first published photographs of cell division appeared in Wilson's 1895 *An Atlas of the Fertilization and Karyokinesis of the Ovum*. He felt that, as an eyewitness account, they provided compelling evidence about cell structures. Such carefully discovered details then led to empirical generalizations, in turn the grounds for new hypotheses.

But the process of turning observation into general principles and theory does not immediately give a perfect reflection of nature. Further, it takes us "beyond the solid ground of fact into a region of more or less doubtful and shifting hypothesis, where the point of view continually changes as we proceed". Yet he held that science is not just speculation, and working hypotheses guide us beyond a collection of facts to an under-



standing of organizing principles and testable theory. Few modern scientists would disagree.

For Wilson, development was not chaotic, as it begins with inherited chromosomes in the nucleus and inherited cytoplasmic structure. Inheritance guides development, which is capable (more or less, because of past evolutionary adaptations) of responding to environmental conditions. The cell binds inheritance and development, and provides the first of two foundations for biology.

Unlike some today, Wilson had a balanced view of life, with neither too much free will in the form of regulatory plasticity, nor too much genetic determinism. Heredity and development must work together, and always in the context of an evolutionary past. Wilson saw the organism as an assemblage of responsive, vital and integrated interdependent parts, and he saw biology as seeking an understanding of all life within one research programme. He was sure that by progressing 'step by step' we can understand the principles driving even the most fundamental life processes. But someone has to integrate the steps and the pieces, as the organism functions and lives as a whole. That is the fine wine of biocomplexity.

Evolution provided Wilson's other foundation for life, and for biology. History is therefore fundamental for biology. All life has evolved

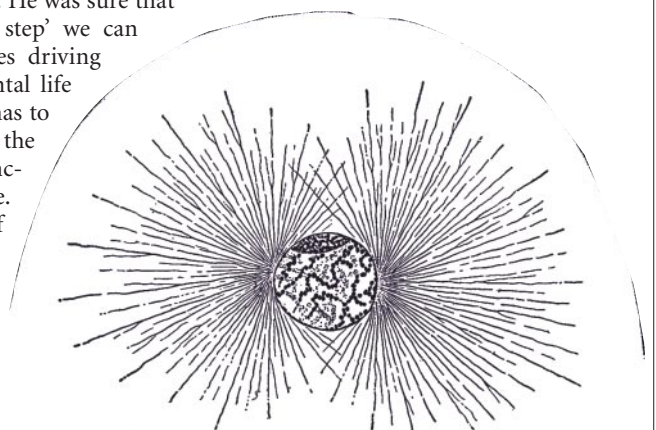
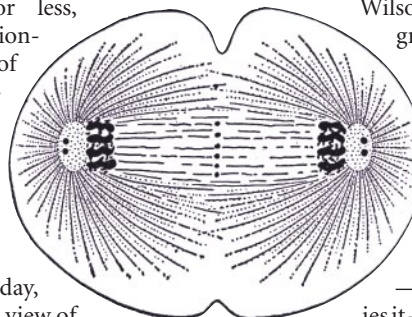
and carries responses to past conditions. For the most part, however, evolutionary processes lie in the past, and experimental biologists need not consider evolution directly when studying cells, development and inheritance. Yet evolution is an essential background condition that scientists ignore at their peril.

Education is important both for researchers and for a wider public understanding. Wilson and his cousin William Sedgwick published the second edition of their *General Biology* in 1896. They sought to promote understanding of the principles of life, starting with fundamental patterns and processes and working up to the complexities of living, functioning, growing, differentiating and reproducing higher organisms. Although beginning with simple examples, Wilson and Sedgwick insisted that life is more than the sum of its mutually interdependent parts and that the organism acts as a living whole.

Wilson outlined a biology programme for his time, and for ours. As every egg contains its past, present and future, Wilson's unified view provides an early 'evo-devo-cellulo' approach to biology. His contributions remind us that life — and the science that studies it — "can best be appreciated from an historical point of view". The

history of science helps us appreciate the old wine, and our growing understanding of biology through the historical events of evolution, inheritance and development provides the new bottles. ■

Jane Maienschein is in the Department of Philosophy, Arizona State University, Tempe, Arizona 85287-2004, USA.



When I was your age

The more things change...

John E. Stith

Ralph slouched against the bed, twisting sections of an antique Rubik's Cube and finding chaos increasing. On the other side of the room, his friend, Terry, sat cross-legged, tossing his baseball into the air and catching it. The window behind him revealed moonlit rooftop spires in Bombay.

"Well?" Ralph asked.

"The parties at Plato's are boring. And you don't even know if you can borrow the scooter."

"I can deal with Dad."

"He's probably still mad about the time you slipped that concoction into your kid brother's food and he lost all his hair."

"It grew back, for Pete's sake. Try me," Ralph said. "You be him."

Terry stopped tossing, twisted to face him. He squared his shoulders, getting into the role. "All right, young man," he said with a deepened voice. "Tell me why I should say yes."

"Mom said it was okay with her if it was okay with you."

"She didn't tell me that. Maybe I should ask her in here."

"This is important, Dad. You say social skills are as vital to getting ahead as what you study. This party lets me develop my social skills."

Terry frowned and said: "When I was your age, we only had ten elements, and most of them were gases."

Ralph snorted. "He'd say: 'When I was your age, I only got to use the scooter for important occasions.'"

"So you'd say —"

"Mom said that would give you two a chance for some quality time, and that you'd know what that meant." Ralph gave the cube another twist. "Got that from Gramps."

"Good one," Terry said, and grinned. "Well, maybe you've got a chance."

"It's like nuking fish in a barrel. See you there at eight my time?"

"Okay, but warn me if you're grounded."

"Deal." Ralph hesitated. "Something else bothering you?"

"Oh, my folks are talking about me getting a job this summer."

Ralph slouched deeper. "Mine too. What is it with parents? They go on about how hard their jobs are and how these are the best years of our lives, and then they want us to get the crappiest jobs we'll probably ever have." After a moment of consideration he added, "I hope."

"Let's talk more at the party. My folks moan that I spend too much time connected."

"Right, see you there." Ralph raised his voice. "Communications. Off." The hologram of Terry and Bombay vanished.

Ralph found Dad in the living room watching the news. A soldier aimed a pulse beam right at him and pulled the trigger, making him flinch. "Dad, I need the scooter tonight for a party at Plato's."

Dad switched off the news and turned to

face him. "All right, young man. Tell me why I should say yes."

So predictable. "Mom said —"

"By the way. I got a call from Mr Hammond at school. Bioengineering class."

"Dad, I can explain —"

"So a virus you created in lab somehow got loose. And it turned everyone's skin blue."

"Well, not everyone's —"

"Oh, right. Somehow you miraculously avoided infection."

"The thing to keep in mind here, Dad, is that everyone is back to normal. It's not like anyone died." Ralph knew this vein wasn't nearing the mother lode.

"When I was your age, I only got to use the family scooter once a month."

"I know, and Gramps could only afford to take it to the end of the block."

"Now don't get smart with —"

"Sorry, Dad. It's just that times change, you know?"

"I still don't see why you need to go to this party."

"You always say how valuable social skills are. This is —" Ralph followed Dad's gaze. Mom was standing in the doorway.

"Mom," Ralph said, "I was just talking to Dad about going to a party. He said that might give you some quality time together tonight."

Ralph could see Dad readying a contradiction, but Mom beat him to the punch with a smile not quite like her normal smile. "Is that so, Honey?" Mom said.

Dad got kind of a dopey smile, too. "If that works for you, Dear."

Checkmate.

Mom just smiled some more. Dad distractedly handed the key-card to Ralph. "Be back by one." He glanced at Mom. "No, make that two."

"Anything you say."

In the scooter, Ralph settled into the deeply cushioned seat. "Autopilot. Plato's, and step on it."

The garage roof slid aside and the scooter rose silently. Once safely over the New Jersey neighbourhood, the craft pitched and yawed into the right attitude, and the gees piled on, pressing Ralph deep into the cushions. The autopilot said, "ETA Plato Crater Spaceport, Luna, twenty hundred hours."

The air outside faded from blue towards black.

Parents were just so old fashioned. He was really going to have to work on them for prom week on Ganymede.

Nebula Award-nominated John E. Stith (www.neverend.com) is the author of *Redshift*, *Rendezvous*, *Manhattan Transfer*, and others.

JACEY



Talk of cannibalism

Jared M. Diamond

Incontrovertible evidence of cannibalism has been found at a 900-year-old site in the southwestern United States. Why do horrified critics deny that many societies have found cannibalism acceptable?

The difficulties of proving the occurrence of cannibalism are illustrated by an experience of mine in New Guinea. On 18 August 1965, while a dozen New Guinea men and I were collecting birds at a remote camp, a man unfamiliar to me arrived and began talking in the Tudawhe language to one of my workers, named Hinobe. The next morning Hinobe departed, claiming that the visitor brought news that his daughter was sick. My other New Guinea friends who heard the conversation later told me the real reason: Hinobe's prospective son-in-law had just died, and Hinobe was expected to join in eating the body.

My friends' detailed account of the eating ceremony matched accounts by Australian patrol officers who arrived unexpectedly in villages when a body was being prepared for consumption. I was not invited to the ceremony, so my experience does not provide first-hand evidence for cannibalism. But archaeologists working in the southwestern United States have now provided compelling evidence. Marlar *et al.*¹, writing on page 74 of this issue, report one episode, and a recent book² offers dozens of examples. Because cannibalism is such a controversial topic, I

report the evidence as an imaginary conversation between claimants and sceptics.

PRO: There are thousands of detailed, independent, mutually consistent accounts in recent times, from some societies but not others, of 'customary cannibalism'. This means the consumption of human flesh (from either deceased relatives or slain enemies) as a non-emergency custom, not just as something that cannibalism-abhorring societies may resort to in an emergency — for example, in 1846 when American pioneers trapped for the winter without food ate each other (the Donner Party).

CON: We consider all of the second-hand accounts to be by gullible or biased observers, and all of the first-hand accounts (none of which is by an anthropologist) to be fabricated. To quote W. Arens³, "There is no satisfactory first-hand account of this act [cannibalism] as a socially approved custom in any part of the world."

PRO: You would consider many accounts convincing if they did not involve a phenomenon that you abhor, or if the observer had been an anthropologist. But there is also abundant archaeological evidence. For example, at many AD 900–1300 Anasazi or

Puebloan archaeological sites in the southwestern United States there are disarticulated bones of dozens of individuals, bearing marks showing that they had been cut up and roasted or boiled² (Fig. 1).

CON: That bone evidence can be interpreted in other ways, such as the slaughter of enemies, scattering of bones by carnivores or scavengers, reburials of people who had died some time ago, or execution of witches.

PRO: Each of these events leaves distinctive macroscopic and microscopic marks on human bones, different from each other and from a characteristic set of six marks on the bones at the putative cannibalism sites². These six marks instead resemble those on the bones of animals eaten at the sites.

CON: But human bodies could have been cut up like animal bodies for reasons other than for eating.

PRO: A hastily abandoned Puebloan site, from AD 1150, has now been studied by Marlar and colleagues¹; at this site, disarticulated, defleshed, cooked bones of seven people were found in a pile or scattered over the floor. A pot at the site contained residues of human myoglobin, a protein confined to human heart and skeletal muscle, as detected by a specific antibody assay. Pots from the same site that were not used for cooking, and cooking pots from other nearby sites, lacked human myoglobin. So someone was cooking human flesh just before the site was abandoned¹.

CON: That supports the cooking of human flesh, but it doesn't prove that the cooks ate the flesh.

PRO: Ash in the fireplace contained dry, unburned human faeces, apparently deposited after the last fire in the fireplace¹. The faeces lacked plant remains, but contained human myoglobin, so the person who deposited them had recently eaten only meat, including human flesh.

CON: But perhaps the myoglobin-specific antibody was not species-specific, and the myoglobin came from an animal.

PRO: The antibody was species-specific, because antibodies that recognize myoglobin of all of the major animal species in Puebloan diets — including deer, antelope, elk, bison, rabbit, rat, turkey, canids and felids — had been removed¹.

CON: But even normal human faeces contain human proteins from shed intestinal



Figure 1 Evidence for the occurrence of cannibalism: the complete collection of human bones from Room 2 of the AD 1150–1200 Anasazi pueblo Houck K, in Arizona. Bones from at least four adults, two teenagers and one child were found. As at many other sites where cannibalism is thought to have occurred, vertebrae are largely missing (because they were crushed to extract marrow), and ends of bone fragments have a rubbed appearance (as a result of having been boiled in pots). Marlar *et al.*¹ have now found further evidence for cannibalism — the presence of human myoglobin protein in cooking pots and in human faeces from an AD 1150 Puebloan site in the southwestern United States. (Photograph reproduced from ref. 2.)

cells. Perhaps the faeces came from an Indian with a bleeding intestine.

PRO: Because myoglobin comes only from heart and skeletal muscle, myoglobin was not detected in 20 'control' faeces from other Puebloan archaeological sites, nor in modern faeces from 29 healthy individuals and 10 patients with blood in their faeces. The myoglobin-specific antibody used does not detect the related blood protein haemoglobin¹.

CON: Okay, you've proved one case of cannibalism, but it's still a rare aberration. Instead of denigrating Indians by looking for proof of bad behaviour, why don't you look for proof of good behaviour?

I ask instead, why is the evidence for cannibalism — which suggests that this practice was once widespread — now so desperately denied? I can think of at least four reasons. First, because Westerners abhor cannibalism, some of us cannot believe that other societies practised (or still practise) it. But many behaviours accepted by one society are abhorred by another. The horror of my New Guinea friends when I described circumcision, US treatment of the elderly, and US funeral customs matched Westerners' horror at cannibalism. Some widespread Western practices are far more destructive than cannibalism. There are good reasons why cannibalism might have been customary in some societies but abhorrent in others⁵.

Second, because Westerners abhor cannibalism, Western missionaries and government officers who encounter a society practising cannibalism immediately forbid it. So it is no surprise that there are few first-hand accounts of cannibalism by twentieth-century Westerners: would you invite someone to watch you doing something if it would get you arrested? Once Western control is established, cannibalism quickly dies out.

Third, those Westerners who obtain evidence of cannibalism are condemned as slandering the non-Western society reported as having practised it — condemned not only by anthropologists offended at perceived insults to 'their' people whom they study, but also by the cannibals' descendants who have absorbed the Western values.

Finally, any society has practices considered acceptable in private but inappropriate to practise in public, in the presence either of anyone else (for example, sex or defecation) or of non-clan members (for example, initiation rites or cannibalism). The abundance of New Guinea babies, my knowledge that babies are conceived by sexual intercourse, and second-hand accounts persuade me that New Guineans practise sex, but I have no first-hand observations of it even after many years there. When I read how vigorously cannibalism's critics deny its existence in the absence of first-hand observations by anthropologists, I find myself imagining a conversation among asexually reproducing extraterrestrials who have conquered the Earth.

PRO: Abundant second-hand evidence convinces me that humans engage in a ritual, too horrible to describe, involving misuse of the urine-producing organs. Our leaders execute humans suspected of performing the ritual.

CON: Have you observed this awful practice yourself?

PRO: No, never. But I have incontrovertible evidence: immunologically detected seminal protein recovered from the vagina of a female human mummy.

CON: This is just one possible interpretation of one aberrant finding. There are other possible interpretations. Why must

you denigrate humans by seeking evidence for bad behaviour, instead of studying good behaviour? ■

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90095-1751, USA.

1. Marlar, R. A., Leonard, B. L., Billman, B. R., Lambert, P. M. & Marlar, J. E. *Nature* **407**, 74–78 (2000).
2. Turner, C. G. II & Turner, J. A. *Man Corn: Cannibalism and Violence in the Prehistoric American Southwest* (Univ. Utah Press, Salt Lake City, 1999).
3. Arens, W. *The Man-eating Myth* (Oxford Univ. Press, 1979).
4. Turner, C. G. II & Turner, J. A. *Man Corn: Cannibalism and Violence in the Prehistoric American Southwest* p. 8 (Univ. Utah Press, Salt Lake City, 1999).
5. Harris, M. *Good to Eat: Riddles of Food and Culture* (Simon and Schuster, New York, 1985).

Chemistry

The smallest fullerene

Martin F. Jarrold

A flat graphite sheet can be made from carbon atoms arranged in hexagons. Inserting pentagons into the sheet will cause it to pucker and curve; adding just twelve pentagons creates enough curvature, in principle, to make the sheet wrap up and link together to form a spherical shell (a fullerene) or even a closed tube. The deciding factor is how the pentagons and hexagons are arranged. The archetypical fullerene, C_{60} , is a sphere with all 12 pentagons evenly distributed over its surface, each one completely surrounded by a ring of

hexagons. This highly symmetric geometry is possible only with precisely 60 atoms. Some other fullerenes have icosahedral (12-fold) symmetry, but it is quite rare¹. C_{180} , with two rings of hexagons around its 12 pentagons, can adopt icosahedral symmetry. So, in principle, could C_{20} . A C_{20} fullerene has no hexagons, just twelve pentagons. This is the smallest fullerene that can exist, although it had never been seen until now. On page 60 of this issue, Prinzbach *et al.*² claim to have prepared not only the C_{20} fullerene, but also the equally elusive 'bowl' isomer of C_{20} (Fig.

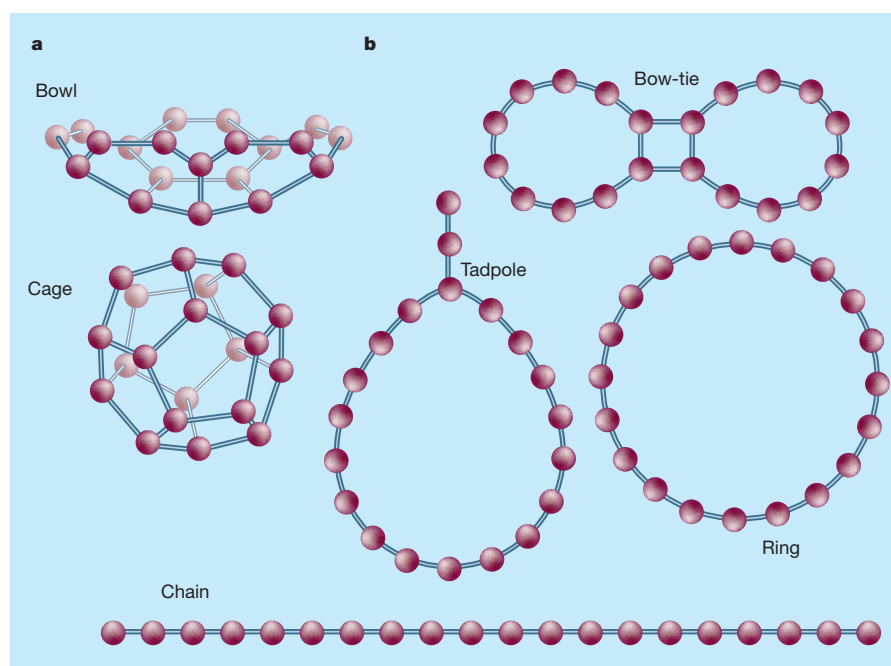


Figure 1 Isomers of C_{20} . a, The fullerene cage and bowl isomers. Prinzbach *et al.*² have created these structures for the first time in minute quantities. Note that the fullerene is expected to be distorted from ideal icosahedral (twelve-fold) symmetry. b, The ring and chain isomers, all of which have been observed previously. Several different forms of the 'tadpole' (a chain attached to a ring) and the 'bow-tie' isomers exist.

1a). The bowl, which has a central pentagon surrounded by five hexagons, has also never been seen before.

The pentagons are the reason why carbon makes such fascinating nanostructures. But they have a dark side: they also cause strain. The best way to handle strain is to distribute the pentagons evenly, and in particular to avoid having two together. The common fullerenes, such as C_{60} , C_{70} , C_{76} , C_{78} and C_{84} , all have isolated pentagons. For fullerenes with less than 60 carbon atoms, however, there are too few atoms to avoid neighbouring pentagons. With arguably one exception, these smaller fullerenes are not stable enough to be prepared in macroscopic amounts: they have been observed only in molecular-beam experiments. In the low-pressure environment of a molecular beam, molecules essentially float around in a vacuum, so it is possible to observe reactive and unstable species that do not survive elsewhere. Early dissociation and 'shrink-wrapping' experiments³ provided indirect evidence that fullerenes could be made in molecular beams with as few as 30 carbon atoms. However, these and other experiments found no evidence for a C_{20} fullerene cage or bowl isomer^{4,5}. It turns out that C_{20} prefers to adopt less compact chain and ring geometries (Fig. 1b), all of which have been observed.

Calculating the relative energies of the different C_{20} isomers is a great theoretical challenge. The most reliable calculations give the bowl the lowest energy, with the ring above the bowl, and the cage having much higher energy still⁶. But although the C_{20} cage is theoretically less stable than the ring and bowl, Prinzbach *et al.*² show that it can be prepared, and survive for at least a short time, if the right precursor is used. Their precursor has the basic carbon skeleton of the fullerene cage capped with hydrogen and bromine atoms. To generate the fullerene, the capping atoms must be removed very gently so that the carbon skeleton does not get hot enough to convert into the lower-energy ring or bowl geometries. A similar scheme was used to prepare the bowl isomer from another precursor. Given that the bowl has the lowest energy, why don't all the rings just convert into bowls? The reason is entropy. There is an entropic Catch 22 in action here. The ring must be heated to overcome the activation barrier and make the bowl. But when the ring is hot enough to convert, the bowl is no longer the favoured (lowest free energy) isomer because the floppier ring has more vibrational entropy.

Determining the structure of C_{20} molecules that are isolated in the gas phase is not easy. The usual structural probes available to chemists, such as NMR and X-ray diffraction, just do not work in this rarefied environment. Prinzbach *et al.* have used a technique called anion-photoelectron spec-

troscopy. Here, C_{20}^- anions are irradiated with a pulsed laser that detaches some of the extra electrons. The kinetic energy of the photo-detached electrons is then determined by precisely clocking their travel time over a predetermined distance. The difference between the photon energy and the energy of the fastest-moving electrons provides a measure of how strongly the extra electron is held by C_{20}^- . If the geometries of the initial anion and the resulting neutral C_{20} are slightly different, removing the electron leaves the neutral C_{20} vibrationally excited. The vibrational frequencies can then be deduced from oscillations in the photo-electron signal.

Prinzbach *et al.* show that the photo-electron spectra of C_{20}^- ions derived from the bowl and fullerene precursors are different from each other, and different from the spectrum for C_{20}^- rings generated by vaporizing graphite with a laser. There is one cautionary note, however. The vibrational frequencies determined for the precursors are not really enough to unambiguously characterize their geometries, because many different C_{20} isomers could have these frequencies. Although it is perhaps unlikely, the C_{20} ions derived from the fullerene precursor could have distorted or partly opened up in some unexpected way. This possibility can be ruled out only by the measurement of other vibrational frequencies.

Now that the bowl isomer has finally been prepared it should be possible to resolve an

important issue about how fullerenes form. There are two conflicting explanations for how fullerenes are made from small carbon fragments: the 'fullerene road' and the 'pentagon road'⁷. The C_{20} bowl is the starting point of the 'pentagon road' process, by which C_{60} fullerenes are thought to grow from bowl to cup to fullerene by adding two carbon atoms at a time. The main challenge to this mechanism was the absence of the bowl isomer, which also precluded any real test of this idea. Such a test should now be possible. In the 'fullerene road' mechanism, small carbon fragments make rings, which then isomerize into fullerenes. This seemingly unlikely isomerization process has already been seen to occur for carbon clusters with more than 40 atoms. The final resolution of this and other issues will hinge on further experiments and on more extensive theoretical studies, both of which will surely be motivated by this latest development. ■

Martin F. Jarrold is in the Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA.

e-mail: mffj@nwu.edu

1. Fowler, P. W. *Chem. Phys. Lett.* **131**, 444–450 (1987).
2. Prinzbach, H. *et al.* *Nature* **407**, 60–63 (2000).
3. O'Brien, S. C., Heath, J. R., Curl, R. F. & Smalley, R. E. *J. Chem. Phys.* **88**, 220–230 (1988).
4. Von Helden, G., Hsu, M. T., Gotts, N. G., Kemper, P. R. & Bowers, M. T. *Chem. Phys. Lett.* **204**, 15–22 (1993).
5. Ott, A. K. *et al.* *J. Chem. Phys.* **109**, 9652–9655 (1998).
6. Grossman, J. C., Mitas, L. & Raghavachari, K. *Phys. Rev. Lett.* **75**, 3870–3873 (1995).
7. Goroff, N. S. *Acc. Chem. Res.* **29**, 77–83 (1996).

Xenotransplantation

New risks, new gains

Jeffrey L. Platt

No field of medicine has stimulated more excitement and controversy than xenotransplantation — the transplanting of animal organs and tissues into humans. The excitement stems from the possibility that transplantation could finally be extended to all patients who need it. The controversies arise from the immunological hurdles to xenotransplantation, and from the possibility that infectious agents might be passed from the animal source to the human recipient and, potentially, more broadly in the human population. In this issue are two papers (first published in electronic form three weeks ago) that speak to this excitement and controversy. On page 86, Polejaeva *et al.*¹ describe how they have cloned pigs — a step towards surmounting the immunological hurdle. Meanwhile, on page 90, van der Laan *et al.*² report that pig pancreatic islet cells, transplanted into mice, can transmit porcine endogenous retroviruses to the mice.

Pigs are ideal sources of xenotransplants because they are available in large numbers and because their organs are similar in size and nature to those of humans. Polejaeva *et al.*¹ cloned pigs by a two-stage approach, which ultimately involved the transfer of nuclei from cultured adult cells to fertilized eggs from which the nuclei had been removed. This approach is similar to that used recently by Onishi *et al.*³, also to clone pigs.

Compared with breeding, cloning is not an efficient way to propagate pigs, but it does have advantages. For example, genes might be added to, or 'knocked out' from, the genome — modifications that are easier to achieve in cultured cells than in whole animals. This technique might provide a way around the rejection of 'foreign' transplanted pig organs by the human immune system. At the moment this problem is sidestepped by damping down the recipient's immune system. But, for example, Polejaeva



100 YEARS AGO

A note in the *Electrician* refers to a curious effect produced by severe thunderstorms upon the glow lamps on the circuits of the Calcutta Electric Supply Co. It appears that immediately following each lightning flash the brightness of the glowing lamps has been observed to increase suddenly, gradually returning to the normal incandescence. This phenomenon has so frequently been observed that the engineers of the company have sought every possible explanation of the curious phenomenon, but have been unable to find any defect in their circuits — which are on the overhead wire system — that might offer an explanation. Indeed, the only conceivable explanation is one which appears so extraordinary that many may find considerable difficulty in accepting it. It is well known that carbon, acting as a coherer in a wireless telegraph apparatus, undergoes the usual sudden decrease in resistance when subjected to electric radiation. It is suggested that the carbon filaments of a glowing lamp may undergo a similar change when exposed to the influence of a tropical thunderstorm in its immediate vicinity. This sudden decrease in the resistance of the filament would, of course, produce a correspondingly rapid increase in its candle-power, after which the gradual self-coherence of the carbon would account for the return of the lamp to its usual incandescence.

From *Nature* 6 September 1900.

50 YEARS AGO

In his presidential address to the tenth All-India Veterinary Congress held last April in Bombay, Dr. S. Datta, director of the Indian Veterinary Research Institute, Mukteswar, appealed for the creation of an Indian veterinary council, which should be a statutory body set up by Act of Parliament, with power to enforce a high standard of veterinary education in Indian veterinary colleges, to suppress quackery in the veterinary profession and to co-ordinate the work of the various State veterinary departments... He deplored the lack of interest shown by the State and the public in the welfare and improvement of farm stock, and rightly urged that veterinarians should not merely wait until disease occurs and then try to remove or prevent it... Quite apart from political or economic considerations, this reorganisation and development is urgently required for the sake of the farm animals themselves.

From *Nature* 9 September 1950.

et al. propose to knock out the gene encoding the enzyme α -1,3-galactosyl transferase, which catalyses the synthesis of a sugar in pigs. This sugar is the major molecule recognized by human antibodies that trigger xenotransplant rejection⁴ (Fig. 1). This possibility seems to bring xenotransplantation closer to the clinic, but may also provoke worries.

Concerns about xenotransplantation stem mainly from the existence of porcine endogenous retroviruses (PERVs), which seem to be present (and harmless, if not useful) in the genome of all pigs. A retrovirus consists of a single RNA strand, enclosed in an envelope of glycoproteins derived in part from the membrane of an infected cell (Fig. 1). After infecting a cell, the RNA part of the retrovirus is copied into DNA, which inserts into the genetic content of the cell and is thus passed on to all of the cell's progeny. The DNA can also direct the formation of new viral particles. The worry is that PERVs might be transmitted first to a human transplant recipient and then more broadly in the population.

Humans have been in close contact with pigs for millennia, with blood products and tissues being exchanged by accident and, recently, by xenotransplantation⁴. Cultured human cells can be infected by PERVs released from cultured pig cells^{2,5}, but studies of 'control' humans and hundreds of patients who have received pig xenotransplants or blood plasma have not revealed even a single case in which PERVs infected human cells in the body^{6,7}. There are three possible explanations for this. First, PERVs may lack the necessary 'equipment' that would enable them to enter and infect human cells *in vivo*. Second, human cells *in vivo* may lack a receptor to which PERVs need to bind in order to gain entry; alternatively, they might not have the machinery needed to allow PERVs to replicate. And third, humans may have natural defences against PERVs that cultured cells do not.

The first of these explanations may now be under threat. Van der Laan *et al.*² have transplanted pig pancreatic islet cells into immunodeficient mice, and found that PERVs infected several tissues — and can perhaps replicate — in these mice. Still uncertain, however, is whether human cells *in vivo* can be infected. Earlier results^{6,7} suggest not, and one might conclude that xenotransplantation is safe. But such a conclusion would be premature, and would ignore the more important lesson to be learnt from van der Laan *et al.*'s work. Rodents carry a variety of viruses and retroviral elements, the genetic material of which can perhaps recombine with that of PERVs. This might give PERVs the equipment needed for them to infect mouse cells — and perhaps even human cells — *in vivo*. The results reported by van der Laan

et al. should serve as a warning: experimental xenotransplants in rodents could also present a risk for humans.

If PERVs have, or might acquire, the ability to infect humans, and so could be thought to pose some risk, then it becomes appropriate to ask whether we could eradicate PERVs from pigs. The method described by Polejaeva *et al.*¹ might offer a means of doing so. But this should not be undertaken unless the risk is certain and substantial, partly because it could have unforeseen effects on the pigs. And, because humans come into contact with pig tissues and blood in other ways (for example, through routine husbandry), the full protection of human society would require the eradication of pigs that were not modified in this way. In my opinion, the best way of determining the risk and consequences of

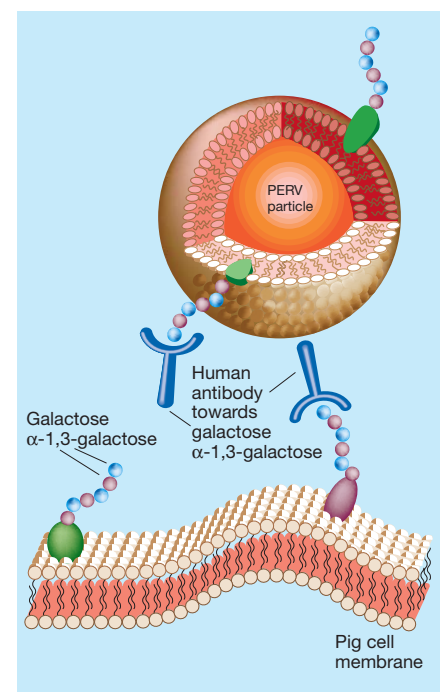


Figure 1 The role of antibodies in the rejection of xenotransplants and in resistance to porcine endogenous retroviruses (PERVs). Most immunocompetent humans have antibodies that recognize galactose α -1,3-galactose, a sugar expressed on the surface of pig cells. The binding of these antibodies to pig cells is thought to trigger the rejection of xenotransplants. Similarly, PERV particles are coated by pig cell membranes, so they also express galactose α -1,3-galactose. So antibodies against galactose α -1,3-galactose might also inactivate PERVs. Some individuals, myself included, who have worked for many years with pigs have low levels of antibodies towards galactose α -1,3-galactose. Such individuals presumably lack the resistance to PERVs conferred by these antibodies. One way of measuring the importance of these antibodies in humans would be to find out whether such individuals are already infected by pig viruses.

human infection by PERVs is to conduct carefully monitored clinical trials of xenotransplants, and to treat experimental xenotransplants in rodents with the care applied to clinical xenotransplants.

Two other possible explanations for the failure so far of PERVs to infect humans are the absence of something — such as a virus receptor — in human cells, or the presence of humans' natural defences. It will be important to find out which of these possibilities is true. If the first is true, PERVs probably pose little or no threat to humans; if the second, PERVs may pose threats that have yet to be considered (Fig. 1).

Human defences against PERVs might consist of T cells, which actively fight infection, or antibodies, which are proteins that bind to foreign substances or particles and help the immune system to eliminate them. One could investigate whether T cells are involved by using van der Laan *et al.*'s experimental set-up to see whether PERVs in pig islet cells can infect mice with functional immune systems. This is important: although the immune systems of patients receiving xenotransplants are, at least at present, suppressed to some degree to avoid the risk of transplant rejection, it is unlikely that they are as deficient as the immune systems of van der Laan *et al.*'s mice.

Resistance to PERVs could also be mediated by antibodies (Fig. 1). Pigs, mice and other lower mammals express α -1,3-galactosyl transferase, which catalyses the synthesis of galactose α -1,3-galactose. This sugar is found on the surface of cells from these animals. Humans and non-human primates do not express this enzyme, but do have antibodies that recognize its sugar product, causing the rejection of xenotransplants⁴. It is tempting to propose, as Polejaeva *et al.*² and others do, that techniques associated with cloning should first be used to knock out α -1,3-galactosyl transferase from pigs. But PERV particles, being wrapped in pig cell membranes, also have galactose α -1,3-galactose on their surfaces. So the same antibodies can also bind to PERVs, neutralizing the virus⁸. Knocking out α -1,3-galactosyl transferase from pigs might render viruses emerging from the xenotransplant less susceptible to inactivation. Whether this genetic modification would endanger humans might also be answered by thoughtful clinical studies. ■

Jeffrey L. Platt is in the Departments of Surgery, Immunology and Pediatrics, Mayo Clinic, Rochester, Minnesota 55905, USA.
e-mail: platt.jeffrey@mayo.edu

1. Polejaeva, I. A. *et al.* *Nature* **407**, 86–90 (2000).
2. van der Laan, L. J. W. *et al.* *Nature* **407**, 90–94 (2000).
3. Onishi, A. *et al.* *Science* **289**, 1188–1190 (2000).
4. Platt, J. L. *Nature* **392** (suppl.), 11–17 (1998).
5. Patience, C., Takeuchi, Y. & Weiss, R. A. *Nature Med.* **3**, 282–286 (1997).
6. Paradis, K. *et al.* *Science* **285**, 1236–1241 (1999).
7. Heneine, W. *et al.* *Lancet* **352**, 695–699 (1998).
8. Rother, R. P. *et al.* *J. Exp. Med.* **182**, 1345–1355 (1995).

Oceanography

Drifters reveal deep circulation

Monika Rhein

In the North Atlantic Ocean, circulation at depths between 500 and 2,000 metres is fed mainly by water from the Labrador Sea (Fig. 1). Here, in winter, surface water cools, becomes denser, sinks to deeper layers, and either circulates within the Labrador basin or passes into the Atlantic. On page 66 of this issue, Lavender *et al.*¹ present the first direct, large-scale measurements of the movement of Labrador Sea water (LSW) in the subpolar North Atlantic. The broader context to their findings is that knowledge of the routes and speed of LSW transport are essential for understanding the ocean's influence on climate and how it varies².

The classical way of measuring flows in the deep ocean is with current meters moored at preselected places. Such single-point data have been used to estimate the direction and volume of water transport in regions where the current speed is swift and the flow is geographically restricted — for instance in the currents 100–200 km wide that run around Greenland and down North America. Together these currents are known as the 'deep western boundary current', which transports LSW and deep water originating north of Iceland to the Southern Hemisphere (Fig. 1).

The current is part of the Atlantic's 'meridional overturning circulation', which is believed to have a large influence on the heat and freshwater budgets of the climate system³.

Single-point measurements are much too expensive and time consuming to be carried out on ocean-wide scales (1,000 km and more). Instead, the large-scale oceanic flows in the Atlantic north of 40° N have mostly been inferred by tracking the characteristic features of the water masses⁴; mean spreading rates are estimated by tracking the changes in these features and relating them to the atmospheric and oceanic changes in the source region. The LSW is relatively cold and fresh, with high levels of anthropogenic tracers such as chlorofluorocarbons; variation in the conditions in which it forms alters these features. With this tracer approach, surprisingly fast spreading of LSW into the Irminger Sea (6 months), the eastern Atlantic (2–6 years) and south into the subtropical Atlantic (6–8 years) has been identified^{5–7}.

Lavender *et al.*¹ have used another, more direct method — the inferred drift of an armada of subsurface floats. Each float drifts at a preselected depth (maximum is 2,000 m) for a few days. It then rises to the surface,

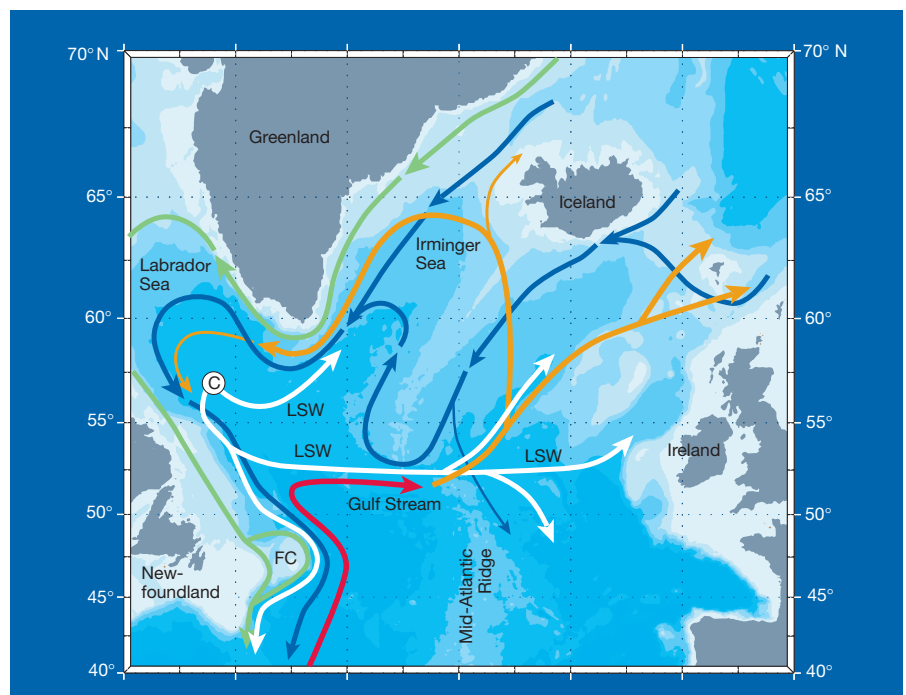


Figure 1 Circulation of the main water masses in the North Atlantic. White shows circulation at intermediate (500–2,000 m) depths; this is the Labrador Sea water (LSW) tracked by Lavender *et al.*¹. Red (warm and saline water) and green (cold and fresh water) indicate circulation in the upper 500 m; the switch from red to orange indicates cooling of warm subtropical water masses in the subpolar North Atlantic. Blue is circulation below 2,000 m. FC, Flemish Cap; C, the region in the Labrador Sea where LSW forms.

measuring the temperature and salinity profile on the way up. On the surface, the float transmits these data and its position to the ARGOS satellite system, before descending and continuing its mission. Mean subsurface drift is calculated from a float's position at two consecutive arrivals at the surface.

The float data, which have been collected since 1994, in general confirm the large-scale circulation in the subpolar North Atlantic and the timescales inferred from tracers. But they also reveal unexpected features — principally, the anticlockwise circulation of LSW in the Labrador Sea itself, in contrast to the strong clockwise circulation along the continental slopes (compare Fig. 1 here with Fig. 2a of the paper on page 66). By this anticyclonic flow, warmer and saltier water might enter the Labrador basin from the northwest corner of the Gulf Stream and increase the density stratification in the LSW formation region (Fig. 3, page 67). More stratified water means that more wintertime cooling is needed to form LSW.

Moreover, the new data show that the main LSW flow path off Flemish Cap, a subsurface plateau east of Newfoundland, resembles that of the Gulf Stream in the upper 500 m (compare Fig. 1 here and Fig. 2a). This indicates that the Gulf Stream participates in the transfer of LSW into the eastern Atlantic.

The question of whether (or how much of) the LSW flows further south in the deep western boundary current is crucial for estimating the effects of LSW formation on the intensity and variability of the Atlantic's meridional overturning circulation. A further surprise in Lavender and colleagues' results is that none of the floats travelled south of 45° N, the meeting point of the southward-flowing boundary current and the northward-flowing Gulf Stream (Fig. 1). Instead, the floats mainly drifted eastwards and remained in the subpolar North Atlantic. On the other hand, we know from tracer observations that LSW is transported on a fast track — presumably the deep western boundary current — into the subtropical Atlantic^{6,7}.

Among other explanations, Lavender *et al.* propose that the LSW route into the subtropical Atlantic might not be in the boundary current, but further east on both sides of the Mid-Atlantic Ridge. Here, however, float coverage is not dense enough to reveal whether this is the case, and the same is true for the tracers. Another reason why the floats did not drift further south in the deep boundary current could be the steep topography at the continental slope near 45° N. This restricts the flow to a narrow band just 30–50 km across, meaning that floats are much less likely to stay in the current. The two methods — tracer evolution and float tracking — complement each other; both will be needed to get the complete picture of oceanic circulation and how it changes.

From the vertical temperature and salinity

profiles in the float data, Lavender *et al.* calculated the thickness of the wintertime, well-mixed surface layer in the Labrador Sea. From these results the precise location and time of LSW formation can be detected, and its intensity estimated. Formation of LSW is governed by wintertime atmospheric conditions and the density stratification in the formation region. The floats can provide stratification data on a basin-wide scale and complement the Labrador Sea data gained by tracer measurements, moored sensors and acoustic tomography.

Finally, a global array of profiling floats is planned in the international Argo programme, which starts this year. The intention is to provide a quantitative description of the evolving state of the upper ocean. This infor-

mation will then be used for initialization and data assimilation in ocean models, and in coupled models of atmosphere–ocean dynamics. The primary emphasis will be on improving our understanding of seasonal to decadal climate variability and predictability. ■

Monika Rhein is at the Institut für Umweltphysik, Abteilung Ozeanographie, Universität Bremen, Kufsteiner Strasse, 28359 Bremen, Germany.
e-mail: mrhein@physik.uni-bremen.de

1. Lavender, K. L., Davis, R. E. & Owens, W. B. *Nature* **407**, 66–69 (2000).
2. Dickson, R. *Nature* **386**, 649–650 (1997).
3. Wunsch, C. *Nature* **405**, 743–744 (2000).
4. Talley, L. D. & McCartney, M. S. *J. Phys. Oceanogr.* **12**, 1189–1205 (1982).
5. Sy, A. *et al.* *Nature* **386**, 675–679 (1997).
6. Molinari, R. L. *et al.* *Geophys. Res. Lett.* **25**, 2249–2252 (1998).
7. Curry, R. G., McCartney, M. S. & Joyce, T. M. *Nature* **391**, 575–577 (1998).

Immunology

RNA editing meets DNA shuffling

Ming Tian and Frederick W. Alt

In its battle against infection, the vertebrate immune system invokes a wide range of immune responses. One tactic is to produce protective proteins known as antibodies (also known as immunoglobulins). Antibodies are secreted by B cells and are encoded by immunoglobulin genes. DNA recombination, in which sequences of DNA are shuffled to create new genetic information, underlies the B cells' ability to generate different antibodies. In fully developed B cells, the immunoglobulin genes have often been shuffled in up to three different ways. The proteins responsible for the first of these 'cut and paste' methods have been identified, but the molecular mechanisms behind the other two processes have been elusive. Now, Revy and co-workers¹ and Muramatsu and colleagues², writing in *Cell*, provide a fascinating and unexpected clue to the solution. The authors offer evidence for the role of a new protein — oddly, an RNA-editing rather than a DNA-recombination enzyme — in both of these processes.

Different antibodies have different functions in the immune system. But they are all made up of a variable region that recognizes foreign substances and a constant region that elicits immune responses to the infection. The first genetic shuffling event that occurs in the immunoglobulin genes is called *V(D)J* recombination. This is used by developing B cells to assemble blocks of DNA into the gene segment that encodes the antibody's variable region³.

During the course of an immune response, mature B cells use a totally different recombination reaction to change the type of constant region in the immunoglobulin 'heavy chain' that is expressed with a

given variable region⁴. This second process is called 'class switch recombination' (CSR).

One of the functions of the heavy chain's constant region is to determine which class of antibody is produced by a cell — for example, IgM, IgG and IgE antibodies all contain different constant regions and have distinct functions. Each constant region is encoded by a different exon (exons are the protein-encoding parts of a gene). After *V(D)J* recombination has occurred in developing B cells, the exons encoding the various constant regions are arrayed sequentially downstream of the *V(D)J* exon (Fig. 1, overleaf). During an infection, when the B cells meet foreign material, the cells 'switch' the class of heavy chain that is expressed, so as to elicit the most effective immune response. This is achieved by CSR, which basically involves replacing the constant-region exon next to the *V(D)J* exon with a downstream exon (Fig. 1).

At the same time, B cells are induced to alter the *V(D)J* exon by 'somatic hypermutation' — the third type of B-cell-specific genetic modification. This process, which is more subtle than CSR, introduces single mutations and small deletions or insertions in the *V(D)J* exon⁵. It allows the production of antibodies with a higher affinity for foreign antigens. So together, CSR and somatic hypermutation tailor antibody production towards particular infections. Until now, these two DNA-shuffling processes were considered to be mechanistically distinct. But the finding that they both rely on the same protein forces a rethink of this idea.

The protein — called activation-induced deaminase (AID) — was discovered last year when it was found to be induced in a B-cell line undergoing CSR⁶. Now, Muramatsu *et*

*al.*² have knocked out the AID gene in mice, and found that these animals are normal except that their B cells completely lack the ability to undergo CSR. The problem seems to lie specifically in CSR, as all known cellular events leading up to this process are normal.

In complementary studies, Revy *et al.*¹ have tracked down the gene that is mutated in a rare form of human immunodeficiency known as hyper-IgM syndrome — a disease in which the B cells can produce only the IgM class of antibody. Remarkably, all the patients with this disease had mutations in the AID gene. The symptoms of this disease¹ are very similar to the characteristics seen in the AID-knockout mice². The specific effect of the AID mutation on CSR is reminiscent of the highly specific inactivation of *V(D)J* recombination in mice that lack the recombination-activating genes³. These genes encode the components of the *V(D)J* recombinase.

Significantly, humans and mice without the AID protein also failed to undergo somatic hypermutation^{1,2} — a fact that came as a surprise. Even more surprising is the nature of AID. Both CSR and somatic hypermutation involve the cutting and pasting of DNA, so one might have expected the proteins needed for these processes would be DNA-recombination enzymes such as nucleases or helicases. But AID belongs to a family of enzymes that 'edit' RNA, changing the base cytosine to uridine by deamination⁶.

Although AID cannot yet be declared the long-sought 'CSR recombinase', it is certainly crucial to CSR. But how can RNA editing be linked to the DNA recombination that clearly occurs during this process? The easiest answer would be that CSR requires a recombinase protein, somatic hypermutation needs a 'mutator', and AID edits the RNA transcripts encoding both. AID is akin to APOBEC-1, a cytosine deaminase that edits the RNA transcript encoding apolipoprotein B. The cytosine-to-uridine change in this transcript creates a 'stop' codon, and so produces a truncated protein with new functions⁷. Perhaps AID works in a similar way. Alternatively, AID might have functions independent of its deaminase activity, and may act directly in DNA recombination.

A final possibility relates to the apparent role of RNA transcripts in CSR². CSR occurs within large regions of 'repetitive' DNA sequences ('switch' regions) that precede each exon encoding a heavy-chain constant region. Transcription through switch regions is induced by stimuli that lead to CSR, and is a prerequisite for this recombination process. Moreover, switch regions transcribed *in vitro* form stable RNA–DNA complexes called 'R loops', which can recruit certain DNA-repair endonucleases⁸. Perhaps AID recognizes and alters the structure of this RNA–DNA complex, enhancing its recognition by recombination or repair systems. Such RNA–DNA complexes have also been

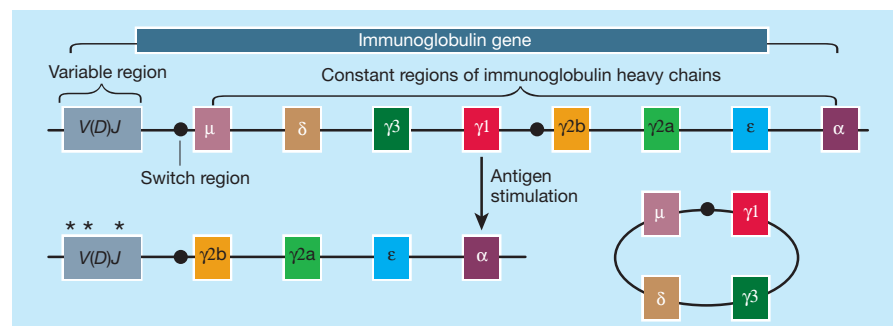


Figure 1 Shuffling antibody genes. The immunoglobulin genes encode antibodies — protective molecules produced by B cells in the vertebrate immune system. The class of antibody produced is determined by the constant-region exon next to the *V(D)J* exon. During an infection, the 'switch' segment of the μ exon recombines with the switch segment of one of the downstream exons (switching with $\gamma 2b$ is shown here). This 'class switch recombination' process results in the replacement of the μ exon by $\gamma 2b$, and the deletion of the intervening sequence. Immunoglobulin G protein is then produced by the cell, instead of immunoglobulin M. Somatic hypermutation also accompanies infection, and occurs (asterisks) within the *V(D)J* exon. The two processes tailor the antibody to the infection. Revy *et al.*¹ and Muramatsu *et al.*² have identified activation-induced deaminase as being crucial to both reactions.

proposed to function in somatic hypermutation⁸; if so, this might explain why AID is needed in both recombination processes.

The discovery of a crucial role for AID in both CSR and somatic hypermutation may well be what is needed to stimulate progress in this field. On another front, switch regions are often involved in the movements of chunks of chromosomes that occur in certain B-cell lymphoma cancers. Conceivably, processes that lead to the unusual properties of switch regions might underlie the instability of related sequences elsewhere in the genome. If so, AID may be important

for more than shuffling antibody genes. ■

Ming Tian and Frederick W. Alt are at the Howard Hughes Medical Institute, The Children's Hospital and Center for Blood Research, 300 Longwood Avenue, Boston, Massachusetts 02119, USA.

e-mail: alt@rascal.med.harvard.edu

1. Revy, P. *et al.* *Cell* **102**, 565–575 (2000).
2. Muramatsu, M. *et al.* *Cell* **102**, 553–563 (2000).
3. Fugmann, S. D. *et al.* *Annu. Rev. Immunol.* **18**, 495–527 (2000).
4. Lansford, R. *et al.* *Molecular Immunology: Frontiers in Molecular Biology* 2nd edn (eds Hames, D. B. & Glover, D. M.) 1–100 (IRL, New York, 1996).
5. Wiesendanger, M. *et al.* *Cell* **94**, 415–418 (1998).
6. Muramatsu, M. *et al.* *J. Biol. Chem.* **274**, 18470–18476 (1999).
7. Navaratnam, N. *et al.* *Cell* **81**, 187–195 (1985).
8. Tian, M. & Alt, F. *J. Biol. Chem.* **275**, 24163–24172 (2000).

Chemical physics

Spinning molecules to bits

Manish Gupta and Richard N. Zare

The speed at which a jet can travel or a centrifuge can spin is limited by rotor disintegration. As the turbine of a jet or the rotor of a centrifuge goes round, it experiences a centrifugal force, which increases in proportion to the square of its rotational velocity. As the rotor speeds up there is some point beyond which the mechanical strength of the rotor cannot withstand the centrifugal force and the rotor breaks. In two papers in *Physical Review Letters*, Corkum and co-workers^{1,2} demonstrate that rotor disintegration will also happen to individual molecules if they are made to spin rapidly enough.

Corkum and colleagues break chlorine molecules (Cl_2) apart by using femtosecond laser technology, which produces short, intense pulses of light whose frequency is made to change with time (to chirp). The oscillations of the electric field of the light pulse, which are perpendicular to the pulse

direction, induce a dipole moment (charge separation) in a molecule. The magnitude of the dipole moment depends on the polarizability of the molecule, which varies with the spatial orientation of the molecule in the electric field. When the electric field oscillations are made to rotate, the dipole moment follows the direction of the electric field, causing the molecule to rotate.

In their optical centrifuge, Corkum and colleagues accelerate the rotating electric field of the light up to terahertz frequencies, forcing chlorine molecules to spin ever more rapidly. In this process, each molecule is excited to increasingly higher rotational states, *J*. As the molecule is driven step by step up the 'ladder' of rotational states, it spins faster and faster, causing the molecule to elongate. Eventually, the centrifugal force exceeds the strength of the molecular bond and the chlorine molecule dissociates into

two chlorine atoms that fly off tangentially to the direction of rotation.

The chlorine atom fragments are then detected by multiphoton ionization, and their velocity measured by time of flight. The results show that many of the atoms have a velocity consistent with centrifugal dissociation. The energy spacing between rotational states varies with J , making it necessary to adjust the frequency of rotation of the light's electric field so that it stays in step with the molecular rotation. This adjustment is accomplished by controlling the chirp of the laser pulse. This experiment is a beautiful example of the power of strong laser fields to manipulate matter. In strong-field optics the strength of the radiation must be comparable to or greater than the electrostatic field of the charged particles in the matter.

The physics behind molecular and rotor disintegration are strikingly similar, despite the different driving forces involved. For a spinning rod, whenever the peripheral velocity (of the outer parts) reaches 2–20% of the speed of sound in the material, the rotor is in danger of flying to bits. A similar calculation can be performed for the chlorine molecule. The peripheral velocity is found from its classical motion in the highest rotational level before dissociation occurs. In the same way as for a solid, the speed of sound is calculated by assuming that it is given by the square root of the ratio of the modulus of elasticity (the second derivative of the potential energy function) to the density of the molecule. For Cl_2 we find that centrifugal dissociation takes place when the peripheral velocity is 13.7% of the speed of sound, which is approximately the same value as for a plastic rod. Remarkably, the same considerations apply to the disintegration of both molecular and macroscopic rotating objects.

When a rotating rod breaks by centrifugation it tears apart at the centre of rotation, which coincides with its centre of mass. This behaviour might seem paradoxical in that this point is stationary, but the force on the rotor at the centre of rotation is the greatest because it is the sum of all the forces from more distant points. The same reasoning applies to spinning chlorine atoms. In this sense, rotational destruction of a molecule could be a new way to select and control bond breaking in polyatomic molecules, a long-sought goal in physical chemistry³. For example, consider a linear polymer of N equally spaced monomer units. In the optical centrifuge built by Corkum and colleagues it would dissociate into polymers of $N/2$ monomer units — break in the middle — in sharp contrast to thermal dissociation (simply heating the molecules up), which yields a wide range of different polymer lengths.

Experiments on the rotational destruction of larger molecules should now be possible

because the method of dissociation used by Corkum and co-workers is so universal, requiring only that the molecule be polarizable. Although polyatomic molecules are more complex and have a much richer energy-level structure, working with larger molecules does have several advantages. In addition to being more polarizable, larger molecules need not be spun as fast as smaller ones to cause dissociation. For example, compare the angular velocity ω_N required to rupture a linear polymer of N equally spaced monomer units with that necessary to break apart a molecule of $N=2$ units. If we assume the same energy for bond cleavage, a tetramer ($N=4$) requires an angular velocity that is about 3.2 times smaller than that for the dimer.

Linear polymers are just one way to use destructive molecular centrifugation. As Corkum and co-workers point out, the preparation of rotationally excited molecules that are not vibrationally excited may offer several advantages. Some possibilities include using them for etching surfaces, promoting new chemical reactions and creating tunable multi-terahertz radiation sources. Let the molecular whirling dervishes begin! ■

Manish Gupta and Richard N. Zare are in the Department of Chemistry, Stanford University, Stanford, California 94305-5080, USA.

e-mail: zare@stanford.edu

1. Villeneuve, D. M. *et al.* *Phys. Rev. Lett.* **85**, 542–545 (2000).
2. Karczmarek, J., Wright, J., Corkum, P. & Ivanov, M. *Phys. Rev. Lett.* **82**, 3420–3423 (2000).
3. Zare, R. N. *Science* **279**, 1875–1879 (1998).

Alzheimer's disease

A partner for presenilin

Dale Schenk

At the core of much of the research into Alzheimer's disease is the mechanism that generates the amyloid- β ($\text{A}\beta$) peptide. This peptide is the building block of the toxic 'plaques' characteristic of brain tissue from patients with Alzheimer's disease. On page 48 of this issue¹, Yu and colleagues describe a new protein — which they call nicastrin after the Italian village of Nicastrò, key to early studies of this disease — that may be involved in generating the $\text{A}\beta$ peptide. Their findings not only advance our understanding of this process and its relationship to Alzheimer's disease, but also provide a new molecular target for the design of drugs to treat this neurodegenerative disorder.

More fundamentally, their work provides an insight into how cells might dispose of membrane proteins that they no longer need.

The $\text{A}\beta$ peptide is generated from a larger precursor, the amyloid precursor protein (APP), by two sequential enzymatic 'activities' called β -secretase and γ -secretase. These activities result in the cleavage of APP in two places to produce $\text{A}\beta$ (Fig. 1). The $\text{A}\beta$ peptide is then released into the brain. APP stretches through the plasma membrane, with domains both inside and outside the cell — the γ -secretase activity makes its cut in part of APP that is within the membrane². This activity also has a more benign and useful role: during development, it might

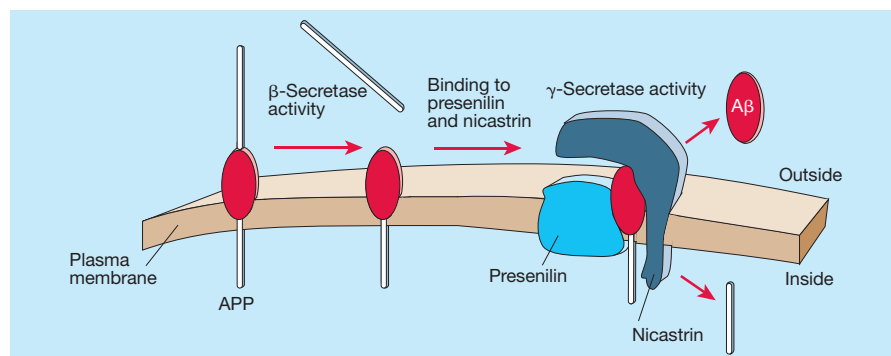


Figure 1 Does a protein complex underlie the γ -secretase activity that generates the amyloid- β ($\text{A}\beta$) peptide? The $\text{A}\beta$ peptide is found in the brain plaques of patients with Alzheimer's disease. It is produced from the amyloid precursor protein (APP) in two steps (left to right). First, APP is cut by a membrane-bound aspartyl protease, called β -secretase, resulting in a secreted segment of APP and a membrane-bound stub. This stub is then processed by the γ -secretase activity, generating the $\text{A}\beta$ peptide. The presenilin proteins are required for this activity, and are likely to be the essential enzymatic component. Yu *et al.*¹ now show that another protein, nicastrin, is also required. It might act to position the APP stub correctly to allow presenilin to cut it at the right place; or it might regulate the activity of the γ -secretase enzyme (possibly presenilin). In the model shown, the large extracellular domain of nicastrin blocks the entry of APP (which is not cleaved by γ -secretase) and other protein substrates containing large amino-terminal domains.

cleave a transmembrane protein called Notch, releasing a fragment that activates the transcription of genes involved in cell-fate determination^{3,4}. But it has been difficult to identify the protein(s) responsible for γ -secretase activity. Part of the difficulty was finding a protein-cleaving enzyme that cuts its targets within a transmembrane domain.

It was discovered in 1995 that a missense mutation (one that results in the insertion of an incorrect amino acid) in a previously unknown protein can lead to an early-onset form of familial Alzheimer's disease⁵. This protein was named presenilin. Since then, a great deal of effort has been devoted to trying to understand how the normal and mutant presenilin proteins can lead to Alzheimer's disease^{6,7}. For example, the presenilin mutants that predispose people to Alzheimer's disease result in APP being cleaved more frequently in a different place to normal, producing a slightly longer and more toxic form of A β .

It now seems likely that presenilin is behind the elusive γ -secretase activity^{8,9}. The data are compelling, but some doubts linger. First, the relative molecular mass of the purified cellular extract that has γ -secretase activity is higher than that of presenilin. And no one has yet been able to show that purified presenilin alone has γ -secretase activity. So, presenilin may not be working alone.

The discovery of nicastrin by Yu *et al.*¹ confirms this supposition. The authors approached the problem of identifying other proteins that may be involved in the γ -secretase activity by purifying large amounts of presenilin from a particular human cell type. They then isolated the proteins that, because they bind to presenilin, were also found in the purified extracts. Two of these proteins were α -catenin and β -catenin, which were already known to bind to presenilin but do not seem to have a role in APP processing. The third protein was a new transmembrane protein of unknown function. Further analysis revealed that this protein, now called nicastrin, binds to both presenilin proteins (presenilins 1 and 2) and interacts with the APP carboxy-terminal 'stub' — the fragment of APP that is produced by the initial, β -secretase-mediated cleavage (Fig. 1). Mutations that alter this interaction also alter the overall processing activity of γ -secretase, either positively or negatively.

To find out whether nicastrin is required for Notch processing, too, Yu *et al.* knocked out the function of nicastrin in the nematode worm *Caenorhabditis elegans*. They found that the offspring of these worms had the same characteristics as those in which the activity of genes in the Notch signalling pathway is reduced. It seems that nicastrin is probably required for γ -secretase activity in this processing reaction as well.

So, presenilin and nicastrin probably form a functional complex involved in the

development of Alzheimer's disease. How might these proteins work together? One possibility (Fig. 1) is that nicastrin binds to the APP stub and aligns it in the correct way relative to presenilin, so that it can be cleaved at just the right position. This would suggest that nicastrin controls the specificity of cleavage but lacks the active site. Another possibility is that nicastrin regulates the cleavage activity, in which case changes in presenilin or nicastrin might independently, or together, have allosteric effects on overall γ -secretase activity and APP turnover. Either way, compounds that interact with either nicastrin or presenilins should effectively alter γ -secretase activity. Indeed, compounds that interact with the presenilins in this way have already been identified^{9,10}. Such compounds might, in the future, be useful in slowing down the progression of Alzheimer's disease.

Does the γ -secretase complex contain other proteins, too? And, on a more fundamental level, is this complex involved in processing and perhaps getting rid of other transmembrane proteins? Such a role would be analogous to that of the 'proteasome', a cellular protein-cleaving machine that

processes some cytoplasmic proteins to their active form, and completely degrades faulty or unneeded cytoplasmic proteins. If the γ -secretase complex treats transmembrane proteins in the same way, a more apt name for it might be 'secretosome' (as suggested by Yu *et al.*; this name would reflect the activity's role in generating secreted peptides such as A β) or 'membrasome'. It seems that the production of the ill-fated A β peptide has led to a far greater understanding of what might — for proteins such as Notch — be a normal cellular process. It is tragic indeed that this process might also contribute to Alzheimer's disease in our old age. ■

Dale Schenk is at Elan Pharmaceuticals, 800 Gateway Boulevard, South San Francisco, California 94080, USA.

e-mail: DSchenk@elanpharma.com

1. Yu, G. *et al.* *Nature* **407**, 48–54 (2000).
2. Selkoe, D. J. *Trends Cell Biol.* **8**, 447–453 (1998).
3. De Strooper, B. *et al.* *Nature* **398**, 518–522 (1999).
4. Schroeter, E. H., Kisslinger, J. A. & Kopan, R. *Nature* **393**, 382–386 (1998).
5. Sherrington, R. *et al.* *Nature* **375**, 754–760 (1995).
6. Haass, C. & De Strooper, B. *Science* **286**, 916–919 (1999).
7. Sisodia, S. S., Kim, S. H. & Thinakaran, G. *Am. J. Hum. Genet.* **65**, 7–12 (1999).
8. Wolfe, M. S. *et al.* *Nature* **398**, 513–517 (1999).
9. Li, Y. M. *et al.* *Nature* **405**, 689–694 (2000).
10. Esler, W. P. *et al.* *Nature Cell Biol.* **2**, 428–434 (2000).

Nanotechnology

Bouncing a C₆₀ ball

Leo Kouwenhoven

Bouncing balls fascinate not only soccer and basketball fans, but also some nanoscientists. On page 57 of this issue, Park *et al.*¹ describe the bouncing of the smallest possible soccer ball, a C₆₀ molecule with a diameter of 0.7 nanometres. Like all classic soccer balls, a C₆₀ molecule consists of 12 pentagons surrounded in total by 20 hexagons². Regardless of the ball's size, the spherical geometry always has the same number of pentagons and hexagons.

Many chemical, electronic and physical properties of C₆₀ have been studied in the 15 years since its discovery². The experiment by Park *et al.*¹ adds mechanical properties to this list. They do with C₆₀ what others do with a ball, bouncing it up and down on a surface. Controlling the motion of nanoscale objects is an important issue in the field of nanotechnology. Whereas in the macroscopic world the transfer of energy from a bouncing tennis ball to a surface is negligible, on the nanometre scale the energy of mobile electrons in the material cannot be ignored. In nanoscale objects, the coupling of electronic and mechanical behaviour can be enough to get a molecule moving, despite the much heavier mass of the molecule compared with the mass of the electron.

The mechanical control of nanoscale

objects will mean smaller, faster and more efficient versions of existing micro-electro-mechanic structures (MEMS), an example of which is the accelerometer that triggers airbags in vehicles. A good example of research into nano-electromechanic structures (NEMS) is provided by Schwab *et al.*³, who made nanoscale bridges out of silicon that can transport heat through specific atomic vibrations. The approach taken by Park *et al.* is to use the natural motion of molecules that are loosely bound to a gold surface.

Park and colleagues have succeeded on two counts. First, they have created a three-electron transistor from a single C₆₀ molecule. As in ordinary silicon field-effect transistors, the voltage on a 'gate' electrode controls the current flowing from the 'source' electrode through the C₆₀ molecule to the 'drain' electrode (Fig. 1a, overleaf). In fact, this is the smallest field-effect transistor ever built. The small size of C₆₀ allows only one electron at a time to hop, or tunnel, on and off the molecule. This means that the device is a so-called single-electron transistor. Second, the single-electron current can both excite and detect the mechanical oscillations of the C₆₀ ball. To understand this electro-mechanical

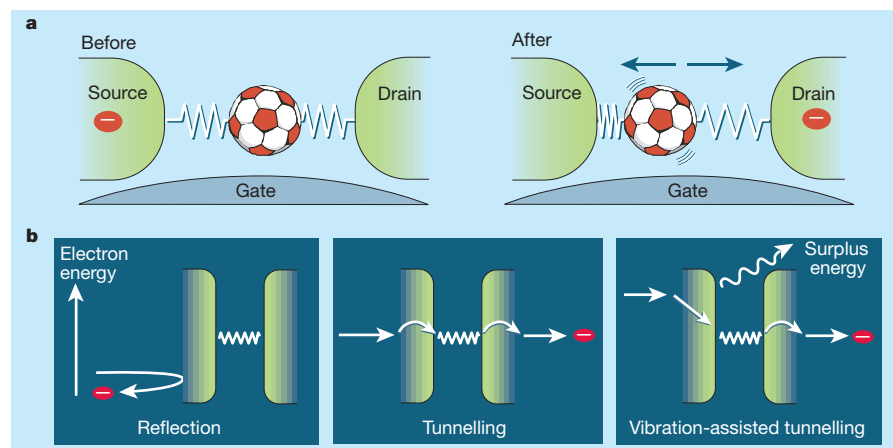


Figure 1 How to build a transistor from a single C_{60} molecule. **a**, The C_{60} transistor can be viewed as a soccer ball bound by two springs to the gold electrodes. First the C_{60} ball is at rest and one electron is at the source electrode. After the electron has tunneled via the C_{60} ball to the drain electrode, it has excited the C_{60} enough to bounce the molecular ball back and forth between the two electrodes. **b**, Possible electron tunnelling processes through the transistor, depending on the energy of the incoming electron. The energy of the electron has to be just right for electron tunnelling: if the energy is too high or too low it will be reflected. But if the C_{60} ball can be made to vibrate it can assist in tunnelling, as Park *et al.*¹ show in their device.

coupling we need to consider the energies that are involved in the different tunnelling processes (Fig. 1b).

To hop on the molecule, an electron has to have the correct energy to occupy a discrete molecular state. Too little energy leads to the electron being reflected, in which case it will not contribute to the current. If the electron has precisely the right amount of energy to occupy the lowest unoccupied molecular state, it can hop on and off, giving rise to electrical current. Too much energy usually also leads to reflection. But in quantum mechanics there exists an extra process by which an electron can tunnel across the molecule, owing to the unavoidable existence of fluctuations even at zero temperature. If the electron has a surplus energy precisely equal to the vibrational energy of C_{60} , then by spontaneous emission of this surplus energy, which starts the C_{60} ball bouncing, it can still hop on and off the molecule. In Park and co-workers' C_{60} device, the applied voltage controls the surplus electron energy. So a sudden current rise at a particular voltage indicates that the C_{60} ball is being made to oscillate.

When bouncing a ball on the ground with your hands, the amplitude and frequency of the bounces are determined mostly by the elasticity of the ball and the forces from gravity and your hands. A similar situation is experienced by the C_{60} ball. The force that makes the molecule stick to the surface of the gold electrodes is the van der Waals interaction. This sticking is not completely rigid. Electrons hopping on the C_{60} ball play the role of the hands, bringing the molecule into motion. But the bounces occur only at particular frequencies, owing to the quantization imposed by quantum

mechanics. When the shape of the C_{60} ball does not deform, the bouncing frequency is about 1 terahertz. If the electrons hit the ball with more energy thereby denting the shape, the bouncing occurs about ten times faster. Park *et al.* found evidence for both types of motion.

In basketball, for regular bounces, the motion of the hand needs to be in phase with the ball's motion. The new experiment by Park *et al.* does not measure or control the phase between the motions of the electrons and the C_{60} molecule. But it has been predicted that, under specific circumstances, every time the C_{60} ball is close to the source electrode an electron might hop on, and when it reaches the drain electrode it would hop off⁴. If during each cycle of the C_{60} oscillation an electron is transferred across, then, because the frequency of the C_{60} bounces is quantized, the electric current also becomes quantized. Electronic devices in which the electro-mechanical motion is strictly coupled in this way could function as 'electron turnstiles' that allow electrons to pass one at a time. Devices in which electrons are under such tight control are being sought to provide a means for measuring electrical current with extreme accuracy⁵.

Leo Kouwenhoven is in the Department of Applied Physics and the ERATO Mesoscopic Correlations Project, Delft University of Technology, PO Box 5046, 2600 GA Delft, The Netherlands.
e-mail: leo@qt.tn.tudelft.nl

1. Park, H. *et al.* *Nature* **407**, 57–60 (2000).
2. Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. (eds) *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
3. Schwab, K., Henriksen, E. A., Worlock, J. M. & Roukes, M. L. *Nature* **404**, 974–977 (2000).
4. Gorelik, L. Y. *et al.* *Phys. Rev. Lett.* **80**, 4526–4529 (1998).
5. Keller, M. W., Eichenberger, A. L., Martinis, J. M. & Zimmerman, N. M. *Science* **285**, 1706–1709 (1999).

Daedalus

Cleaning the cleaner

The traditional cleaning cloth or wiping rag is a universal accessory in every kitchen, workshop and laboratory. But it is ecologically very unsound. The cloth either has to be cleaned in its turn, or must be thrown away, adding to the growing pile of organic waste. Seeking a better technology, Daedalus has been inspired by a kitchen household hint. To clean a dirty kitchen table, wipe it down with a kitten; then hand the kitten back to its mother for grooming.

This primitive biological 'disposal at source' can clearly be improved. DREADCO's biologists are devising a cleaning-rag bearing a carefully formulated mixed bacterial ecology. Some of the organisms degrade hydrocarbons, some hydrolyse proteins, and others split fats. The DREADCO 'Dirt Eater' will be made by dipping a thick inert fabric impregnated with polymerization catalyst into a bacterial culture containing suitable monomers. It will acquire a thin polymer coating loaded with trapped bacteria. A lightly crosslinked alkyd resin should resist bacterial attack while allowing water and organic molecules to diffuse readily through it. Bacteria are immortal, but trapped in the polymer they will have no room to divide. They will also be unable to escape to contaminate objects cleaned by the cloth. After each use, the Dirt Eater will slowly 'digest' the dirt it has picked up, turning it to gases or simple water-soluble molecules. A hygroscopic component in the cloth will prevent it drying out during its digestion period. An active kitchen or workshop would use a set of Dirt Eaters in rotation.

Dirt Eaters will rapidly replace traditional kitchen rags, domestic mops and bathroom flannels. They could even transform medical practice. A self-cleaning wound dressing would save the repeated work and distress of changing such dressings. For this service, however, even a small escape of bacteria could not be tolerated. Daedalus may have to devise a sterile Dirt Eater containing not trapped bacteria, but supported enzymes from them.

Yet even the most voracious bacterial or enzyme system will be unable to digest all possible contaminants. The Dirt Eater may also need a layer of that powerful photo-oxidation catalyst titanium dioxide. When it shows signs of indigestion, it could be laid out in the daylight for a while. Biologically resistant contamination would be mineralized, freeing the Dirt Eater for re-use.

David Jones

Tyrosine phosphorylation in plant bending

Puckering in a ticklish plant is controlled by dephosphorylation of its actin.

The phosphorylation of the amino acid tyrosine in animal proteins acts as an on-off switch in numerous pathways that regulate growth, differentiation and oncogenesis^{1,2}. In the unicellular slime mould *Dictyostelium*, tyrosine-phosphorylation of actin controls cell-shape changes and spore activity^{3,4}, but the significance of protein-tyrosine phosphorylation in plants is unknown⁵. Here we show that actin in the contact-sensitive plant *Mimosa pudica* L. is heavily tyrosine-phosphorylated and that changes in the extent of phosphorylation correlate with the degree of bending of the plant's petioles. We propose that tyrosine-phosphorylation of actin controls movements in plants.

M. pudica closes its leaves and points its petiole downwards when touched (Fig. 1a,b). Petiole bending is caused by a rapid shrinking of the lower side of motor organs called the main pulvinus⁶⁻⁸ (Fig. 1c,d). Cytochalasin B and phalloidin, which both have inhibitory effects on the actin cytoskeleton, inhibit the bending of the petiole, suggesting that actin is involved in bending⁹. We found that actin filaments in the motor cells at the lower side of the pulvinus, but not at the upper side, become more peripheral after bending (data not shown), although it is not known what molecular mechanism regulates these changes.

We carried out two-dimensional polyacrylamide gel electrophoresis on total proteins from the whole pulvinus before and after bending, and immunoblotted them with the anti-actin antibody N350 and reprobed with the anti-phosphotyrosine antibody PY20. Densitometry indicated that, before movement, about 80% of actin

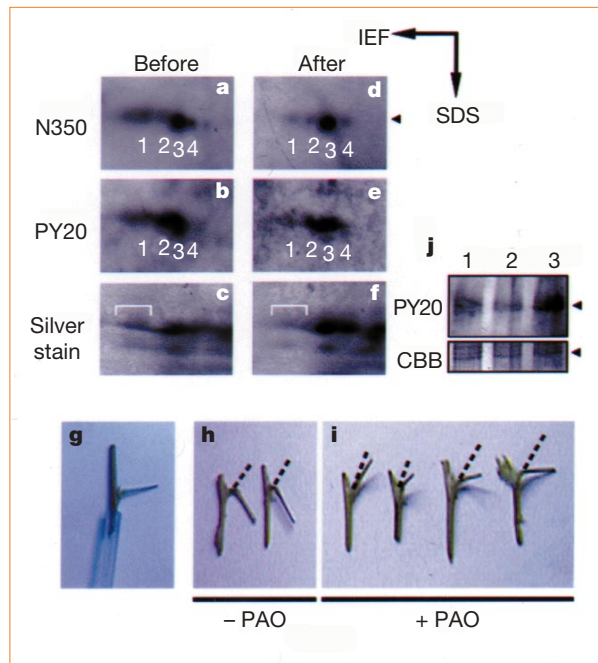


Figure 2 Correlation between the amount of actin tyrosine-phosphorylation and bowing in *M. pudica*. **a-f**, Total proteins from the whole pulvinus before (**a-c**) and after (**d-f**) petiole bending were immunoblotted with the anti-actin antibody N350 (**a,d**) or anti-phosphotyrosine antibody PY20 (**b,e**), or were silver-stained (**c,f**) after two-dimensional polyacrylamide gel electrophoresis (IEF; isoelectric focusing; SDS, denaturing polyacrylamide gel). Four actin spots were detected (numbered): spots 1-3 are tyrosine-phosphorylated, spot 4 is not. **g-j**, Phosphorylation of actin from petioles before the plant is touched (**g** and **j**, lane 1), and after touching (**h,i**) in the absence (**h** and **j**, lane 2) or presence (**i** and **j**, lane 3) of phenylarsine oxide (PAO), a protein-tyrosine phosphatase inhibitor. Dotted lines indicate the original position of the petiole; arrowheads, actin bands.

molecules in the pulvinus were tyrosine-phosphorylated (Fig. 2a,b). After movement, the total amount of actin did not change, but immunoblots with N350 and PY20 on one-dimensional polyacrylamide gel electrophoresis showed that the amount of tyrosine phosphorylation in actin had dropped (data not shown).

Consistent with this decrease in tyrosine phosphorylation, we found that phosphorylated actin in spot 1, the most acidic isoform, disappeared (Fig. 2d-f) — probably because it had been dephosphorylated and shifted to a more basic isoform as a result of bending. In contrast, actin molecules in the stem, which does not move, were heavily tyrosine-phosphorylated, and the level of phosphorylation did not change during movement (data not shown). Changes in the levels of tyrosine phosphorylation of actin in the pulvinus therefore coincide with petiole bending.

In *Dictyostelium* spores, a decrease in actin phosphorylation is a prerequisite for recovering the dynamic actin cytoskeleton involved in spore germination. A specific inhibitor of protein-tyrosine phosphatases, phenylarsine oxide, inhibits the dephosphorylation of actin and the restoration of cell motility during germination⁴. We investigated the effects of phenylarsine oxide on *M. pudica*. When touched, the average bending angle of petioles treated with phenylarsine oxide (Fig. 2g-i) was smaller (37.3 ± 20.24 degrees; $n = 25$) than that of controls (99.0 ± 4.24 degrees). There was no

decrease in the amount of actin phosphorylation following stimulation for the pulvinus exposed to phenylarsine oxide (Fig. 2j).

The bending of *M. pudica* is therefore correlated with reduced actin phosphorylation in the pulvinus. We propose that actin phosphorylation is essential for petiole bending resulting from actin cytoskeletal alterations. Actin may also be phosphorylated during the opening and closing of guard cells, as actin is dynamically arranged in these cells¹⁰.

Kazuhisa Kameyama*, **Yoshiro Kishi†**, **Masanori Yoshimura***, **Nobuyuki Kanzawa***, **Masazumi Sameshima†**, **Takahide Tsuchiya***

*Department of Chemistry, Faculty of Science and Technology, Sophia University, 7-1 Kioicho, Chiyoda-ku, Tokyo 102-8554, Japan

e-mail: t-tsuchi@hoffman.cc.sophia.ac.jp

†Department of Cell Biology, Tokyo Metropolitan Institute of Medical Science, Tokyo Metropolitan Organization for Medical Research, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 171-0021, Japan

1. Sun, H. & Tonks, N. K. *Trends Biochem. Sci.* **19**, 480-485 (1994).
2. Boyle, W. J. *Curr. Opin. Oncol.* **4**, 156-162 (1992).
3. Howard, P. K., Sefton, B. M. & Firtel, R. A. *Science* **259**, 241-244 (1993).
4. Kishi, Y., Clements, C., Mahadeo, D. C., Cotter, D. A. & Sameshima, M. *J. Cell Sci.* **111**, 2923-2932 (1998).
5. Barizza, E., Lo Schiavo, F., Terzi, M. & Filippini, F. *FEBS Lett.* **447**, 191-194 (1999).
6. Datta, M. & Dutt, A. K. *Nature* **179**, 253-254 (1957).
7. Toriyama, H. & Sato, S. *Cytologia* **36**, 359-375 (1971).
8. Tamiya, T. et al. *J. Biochem.* **104**, 5-8 (1988).
9. Fleurat-Lessard, P., Roblin, G., Bonmort, J. & Besse, C. *J. Exp. Bot.* **39**, 209-221 (1988).
10. Eun, S. O. & Lee, Y. *Plant Physiol.* **115**, 1491-1498 (1997).

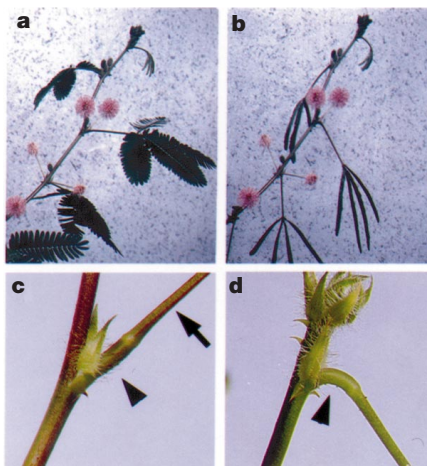


Figure 1 *Mimosa pudica* L. responding to stimulation by touch. **a,c**, Plant before touching and **b,d**, after touching; **c,d**, shows the main pulvinus (arrowheads), with a petiole (arrow)

Structure resolution

Imaging light atoms by X-ray holography

We demonstrate here the imaging of light atoms by hard X-ray holography. Using intense synchrotron radiation for the excitation, the oxygen and nickel atoms in a nickel oxide sample are imaged in three dimensions with high resolution inside a volume of about 1,000 cubic ångströms. These remarkable characteristics mean that X-ray holography bridges the gap between diffraction, which relies on long-range order, and extended X-ray absorption fine structure, which gives information on the local environment of atoms.

For atomic resolution in solids by hard X-ray holography, selected atoms in the sample serve as reference points and their local environments are holographically imaged^{1–3}. After holographic reconstruction, the intensity of the atoms is proportional to the square of their atomic number (z) and also to r^{-2} , where r is the distance from the central atom. To see an atom with a low value of z or at a large distance from the reference, the holographic oscillations therefore have to be measured with great precision.

To achieve this precision, we have used synchrotron radiation and a sophisticated experimental set-up^{4,5}, carrying out our measurements at the two undulator beam-lines ID 32 and ID 22 of the European Synchrotron Radiation Facility. To avoid the twin images inherent in the traditional holographic type of reconstruction, which can distort the three-dimensional real-space image⁶, we collected holograms at eight different energies (from 16.4 to 19.9 keV, in 0.5-keV steps)^{3,7}.

For data handling⁴, we incorporated background subtraction, deconvolution of the contribution of the complementary hologram, extension of the hologram to the full sphere by using the measured symmetries⁵, low-pass filtering⁸ and, finally, multiple energy reconstruction by the Helmholtz–Kirchhoff transformation modified according to Barton⁷.

The results of the measurements and analysis are shown in Fig. 1. The hologram measured at 17.9 keV incident energy is depicted after background subtraction (Fig. 1, top left): the different symmetry points of the hologram and the standing-wave line pattern (the narrow dark and bright lines) can be readily identified. We used these to extend the hologram⁵ to those areas that could not be measured (shaded in grey) (another seven holograms have similar features and comparable quality).

In the final three-dimensional real-space image, the oxygen atoms are evident (Fig. 1, bottom left) (only a limited part of the reconstruction is shown for clarity). The

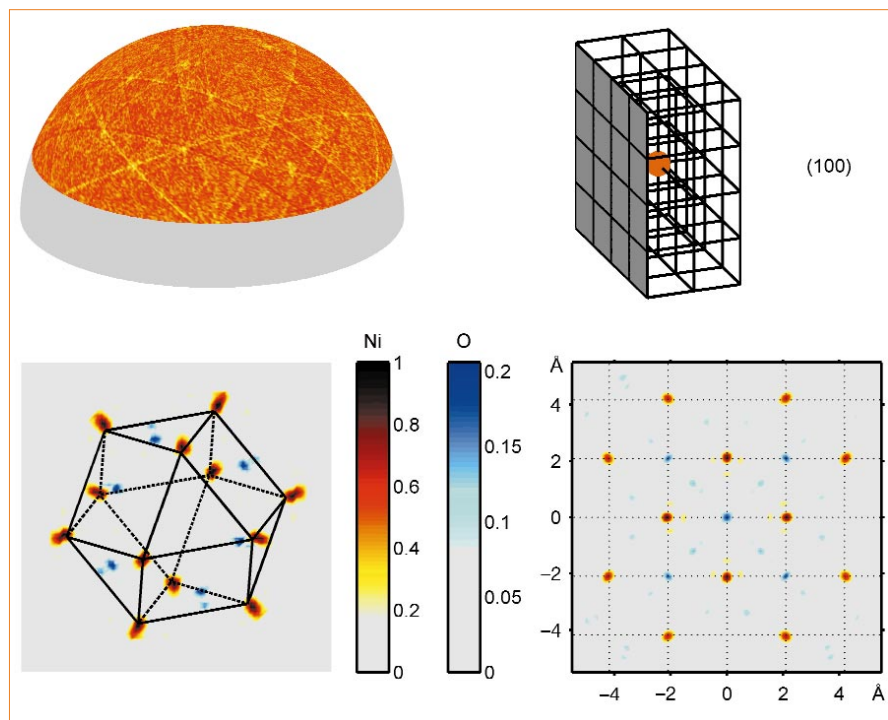


Figure 1 Holographic imaging of atoms in a nickel oxide crystal. Top left, the hologram of NiO is shown in three dimensions on a spherical surface that extends to the half sphere. Shaded grey area indicates the part of the full holographic information that could not be measured. Bottom left, three-dimensional real-space image of atoms within a sphere of radius 6 Å. Oxygen atoms are blue and nickel atoms are red. Bottom right, (100) atomic plane. Atoms can be reconstructed over large distances with isotropic resolution. Top right, for clarity it is also shown how the reconstructed plane relates to the lattice and to the central atom. The colour scale is proportional to the holographic reconstructed intensity.

precise measurement of these holograms allowed us not only to reconstruct the oxygen atoms, but also to image the much more distant nickel atoms up to the seventh coordination shell. This means that there are about 150 atoms in the volume imaged holographically. The (100) atomic plane is shown with distant atoms in Fig. 1, bottom right, illustrating the high and isotropic resolution of the measurement.

Holographic imaging might be extended to systems in which long-range translation periodicity is not present and where X-ray diffraction or extended X-ray absorption fine structure cannot be used efficiently, such as for quasicrystals or single molecules. In quasicrystals, we could see atomic decorations directly. A combination of fourth-generation free-electron laser-type X-ray sources with holographic imaging and reconstruction should lead to the structural determination of single molecules, viruses and other minute systems that cannot be crystallized.

Miklós Tegze*, **Gyula Faigel***,
Stefano Marchesini†, **Michel Belakhovsky†**,
Olivier Ulrich†

*Research Institute for Solid State Physics and Optics, 1525 Budapest, PO Box 49, Hungary

†Service de Physique des Matériaux et Microstructures, Département de Recherche Fondamentale sur la Matière Condensée, CEA Grenoble, 17 rue des Martyrs, 38054 Grenoble, France

1. Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Applications* (eds Attwood, D. T. & Boker, J.) 361–367 (AIP Conf. Proc. no. 147, American Institute of Physics, New York, 1986).
2. Tegze, M. & Faigel, G. *Nature* **380**, 49–51 (1996).
3. Gog, T. et al. *Phys. Rev. Lett.* **76**, 3132–3135 (1996).
4. Faigel, G. & Tegze, M. *Rep. Progr. Phys.* **62**, 355–392 (1999).
5. Tegze, M., Faigel, G., Marchesini, S., Belakhovsky, M. & Chumakov, A. I. *Phys. Rev. Lett.* **82**, 4847–4851 (1999).
6. Len, P. M., Thevuthasan, S., Fadley, C. S., Kaduwela, A. P. & Van Hove, M. A. *Phys. Rev. B* **50**, 11275–11278 (1994).
7. Barton, J. J. *Phys. Rev. Lett.* **67**, 3106–3109 (1991).
8. Tegze, M. & Faigel, G. *Europhys. Lett.* **16**, 41–46 (1991).

Neurophysiology

Good memories of bad events in infancy

If a helpless newborn infant does not form an attachment to its care-giver, even an abusive one, its chances of survival diminish, so evolution should strongly favour attachment by the infant, regardless of the quality of care-giving¹. As a part of the brain called the amygdala is critical for learned fear in adult animals^{2–4}, we investigated whether the development of learned avoidance behaviour could be delayed by late maturation of amygdala function. We found that very young rat pups exposed to various odours associated with shock treatment learn an approach response to that odour, whereas older pups learn odour

avoidance. We show that the origin and development of learned odour-avoidance behaviour is associated with enhanced neural responses in the amygdala during odour-shock conditioning.

Attachment in infant rats is learned predominantly postnatally, relying mainly on the olfactory system — learned odour cues are necessary for successful nipple attachment and for orientation to the mother and litter⁵. As a model for early attachment to an abusive care-giver, we used odour-shock pairings in neonatal rats (fear-conditioning paradigm^{6,7}) that result in odour approach by very young rats but an odour aversion in older rats. Pups were trained in a classical-conditioning paradigm in which odour is paired with shock, or they were exposed only to the odour. Four hours after training, pups were tested for relative odour preference in a two-odour choice test.

Figure 1a shows that the conditioning of approach to odours paired with aversive shock occurs during a period that ends

before postnatal day (PN) 10. Pups trained after this sensitive period learn to avoid odours paired with shock (analysis of variance (ANOVA), conditioning group $\times$ postnatal age, $F_{2,54} = 9.29$, $P < 0.001$; Fisher post-hoc tests indicate that the PN9 paired pups are significantly different from the PN10 and PN11 pups, $P < 0.01$). This striking difference in the learned behavioural responses of PN9 and older pups occurs despite their similar responses to shock (vocalization, vigorous physical response, wall climbing^{6,7}; Fig. 1c, d). Furthermore, both PN9 and PN10 pups try to escape from the shock (results not shown).

Once acquired, a learned odour-approach response to odours paired with shock before PN10 continues to be expressed as an approach even in older pups that are capable of learning an aversion (good memory of a bad event). We trained pups with odour-shock pairings on PN9 and tested them after either 4 or 24 hours (on PN10). As shown in Fig. 1b, pups trained on PN9 and tested either on PN9 or

PN10 showed a relative odour preference for the odour paired with shock (ANOVA, main effect of conditioning group, $F_{1,33} = 17.94$, $P < 0.001$; Fisher post-hoc tests indicate that each paired group is significantly different from each control group, $P < 0.01$).

To determine the role of the amygdala in the differential response to odour-shock training, we injected PN8 and PN12 pups with ¹⁴C-2-deoxyglucose and subjected them to a 45-min odour-shock or control classical-conditioning procedure. Amygdala uptake of labelled 2-deoxyglucose was expressed relative to uptake in the corpus callosum, which did not vary across conditioning groups. The relative uptake by the amygdala was higher in PN8 pups than in PN12 pups, but did not vary with conditioning group in PN8 pups. However, it was significantly enhanced in PN12 pups trained in odour-shock pairings compared with control pups (ANOVA, training group $\times$ age interaction, $F_{2,35} = 3.37$, $P < 0.05$; post-hoc Fisher tests revealed PN12 paired pups were significantly different from both PN12 control groups, $P < 0.05$).

Our results show that the appearance of a learned odour aversion using this conditioning paradigm originates with the initial appearance of an aversive conditioning-induced enhancement of neural activity within the amygdala. We cannot infer a causal connection between amygdala development and ontogeny of avoidance conditioning, however. In fact, illness-induced odour aversions and conditioned taste aversions can be induced in pups during their first postnatal week, although illness-induced aversions may involve a different mechanism from fear conditioning⁷⁻⁹. Nonetheless, our results indicate that a unique behavioural¹⁰ and neural response to aversive conditioning may exist in helpless newborns during a time when aversion to a care-giver could be fatal.

Regina M. Sullivan, Margo Landers, Brian Yeaman, Donald A. Wilson

Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019, USA

email: rsullivan@ou.edu

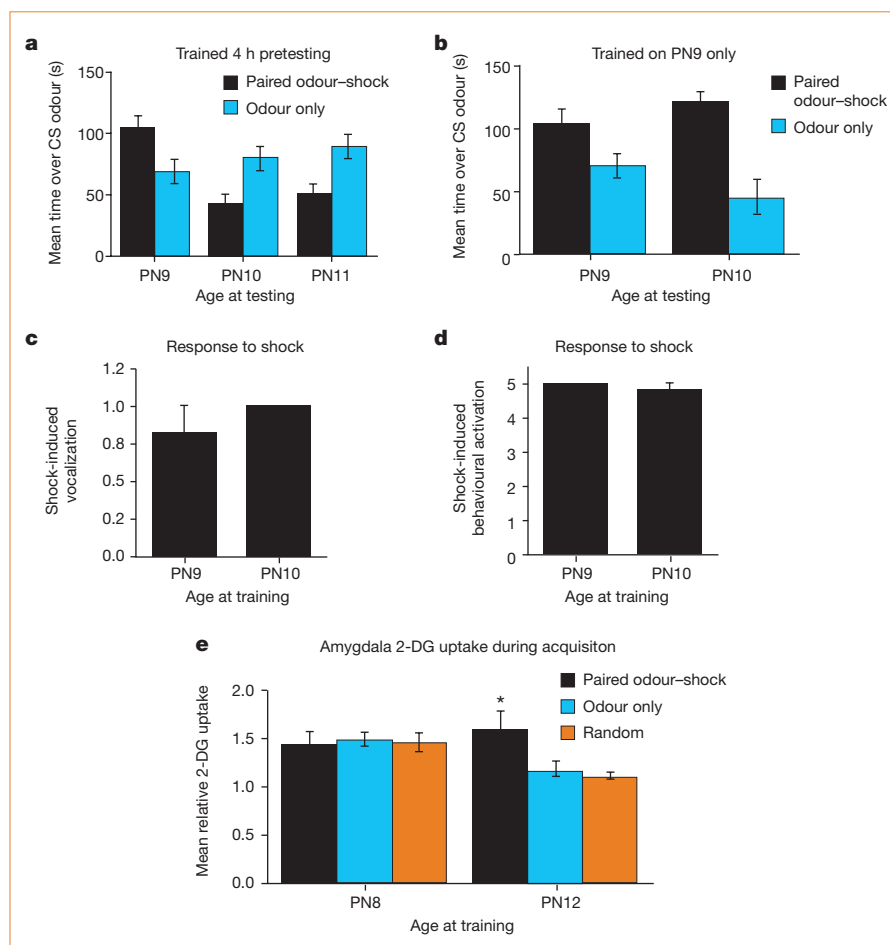


Figure 1 Ontogeny of learned odour aversions corresponds to the ontogeny of amygdala activation. **a**, Postnatal-day (PN)-9 pups trained with paired odour-shock demonstrated a subsequent odour preference compared with PN9 controls. Older pups showed learned odour aversions ($n = 8-12$ per group). **b**, Pups trained on PN9 continued to show an odour preference for an odour paired with shock, even at an age when they could learn an odour aversion ($n = 8$). **c,d**, PN9 pups find shock as aversive as PN10 pups (as measured by shock-induced vocalizations, in **c**, and behavioural activation in **d**; $n = 5-6$). **e**, Association of odour and shock produces enhanced neural activity within the amygdala of PN12 pups compared to controls, whereas the same training does not differentially affect amygdala activity in PN8 pups ($n = 6-8$). CS, conditioning stimulus; 2-DG, 2-deoxyglucose.

1. Bowlby, J. *Attachment* Vol. 1 (Basic Books, New York, 1969).
2. Fanselow, M. S. & LeDoux, J. E. *Neuron* **23**, 229-232 (1999).
3. Davis, M. in *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction* (ed. Aggleton, J. P.) 255-305 (Wiley-Liss, New York, 1992).
4. Aggleton, J. P. *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction* (Wiley-Liss, New York, 1992).
5. Wilson, D. A. & Sullivan, R. M. *Behav. Neurobiol.* **61**, 1-18 (1994).
6. Camp, L. L. & Rudy, J. *Dev. Psychobiol.* **22**, 25-42 (1989).
7. Haroutunian, V. & Campbell, B. A. *Science* **205**, 927-929 (1979).
8. Chambers, K. C. *Annu. Rev. Neurosci.* **13**, 373-385 (1990).
9. Kucharski, D. & Spear, N. E. *Dev. Psychobiol.* **17**, 465-479 (1984).
10. Hunt, P. S. & Campbell, B. A. in *Learning, Motivation and Cognition* (eds Bouton, M. E. & Fanselow, M. S.) 53-74 (American Psychological Association, Washington DC, 1997).

Immunology

Investigating T-cell memory

Memory is a long-recognized, crucial and poorly understood property of adaptive immunity. Jacob and Baltimore¹ have designed an elegant genetic approach to marking memory T cells and their precursors irreversibly and have obtained intriguing results^{1,2}. I suggest that the particular enzyme (human placental alkaline phosphatase) chosen for irreversible marking in postnatal life contributes to certain features of this system and perturbs the homeostatic mechanisms of T-cell activation and death.

Jacob and Baltimore used Cre-recombinase-mediated genetic recombination, driven by the granzyme-B promoter, to obtain irreversible and selective surface expression of placental alkaline phosphatase (PLAP; driven by the CD2 promoter) in T cells activated by antigen in adult life. In principle, this system should result in the irreversible marking of activated T cells and enhance the visualization, purification and fate mapping of activated T cells and memory T cells. As such, it should represent an invaluable tool for analysing T-cell memory, a subject central to the pathophysiology of immunity and vaccine development².

However, this genetic marking approach yielded some unexpected and unexplained results^{1,2}. For instance, in unimmunized mice, none of the T cells expressing high levels of CD44, which have presumably responded to environmental antigens, was positive for the PLAP marker; upon exposure to antigen, only a fraction of activated T cells became PLAP-positive. During a response to lymphocytic choriomeningitis virus (LCMV), only about 10% of LCMV-specific T cells became PLAP-positive.

Different forms of alkaline phosphatase have been described³. Human PLAP shows only limited homology to the mouse isoforms, sharing 55% identity with mouse PLAP for example⁴. Antibodies against human PLAP have thus been generated easily in various animal species, including mice³, so expression of human PLAP in T cells activated in adult life may result in a cellular and/or humoral anti-human-PLAP immune response.

An anti-human-PLAP response could explain some of the unusual features of the results obtained from irreversibly marking activated T cells with this human enzyme. For instance, selective pressure may favour weakly expressing T-cell clones. In general, the activation of immune responses directed against activated T cells may perturb the system in an unphysiological way.

These considerations caution against the overinterpretation of results derived from

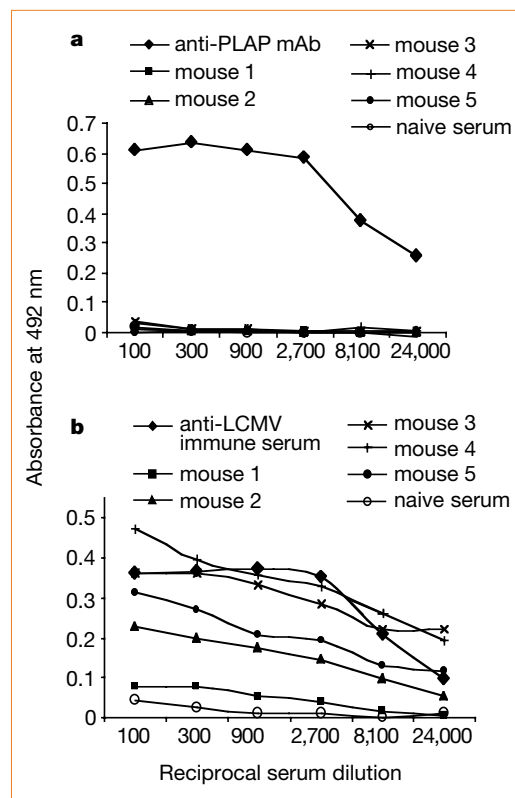


Figure 1 Antibody production by doubly transgenic mice infected with lymphocytic choriomeningitis virus (LCMV). **a,b**, Mice do not produce antibodies against placental alkaline phosphatase (PLAP) (**a**), but are able to mount a good antibody response against LCMV (**b**). For enzyme-linked immunosorbent assays, 96-well flat-bottom Immulon plates (Nunc) were coated with purified PLAP protein (Anawa Trading SA, Zurich; 80 ng per well) for 2 h at 37 °C (**a**) or with a lysate from LCMV-infected BHK-21 cells (**b**). Plates were blocked with 4% non-fat skimmed milk and 0.1% Tween 20, and duplicate dilutions of serum from GranzymeB-Cre × CD2-STOP-PLAP doubly transgenic mice (infected 9 d previously with LCMV Armstrong) were allowed to bind for 90 min. A monoclonal antibody against human PLAP (clone 8B6; Sigma) and LCMV immune serum were used as positive controls for detecting PLAP and LCMV, respectively. After extensive washing, plates were incubated with goat anti-mouse IgG conjugated to horseradish peroxidase (Sigma) for 90 min. After washing, the substrate *o*-phenylenediamine (Sigma) was added and the colour developed was quantified 30–40 min later on a Spectramax 340 plate reader (Molecular Devices) at 492 nm (ref. 3). Background absorbance from wells with no antigen was subtracted from antigen-containing wells at each dilution. The absorbance at 492 nm is plotted against the reciprocal of the serum dilution.

the indelible marking of activated T cells with human PLAP. To be informative, the same genetic approach should take advantage of non-immunogenic tracers.

Alberto Mantovani

Istituto di Ricerche Farmacologiche 'Mario Negri', Via Eritrea 62, 20157 Milan, and Department of Biotechnology, Section of General Pathology, University of Brescia, via Valsabbina 19, 25123 Brescia, Italy
e-mail: mantovani@irfmm.mnagri.it

- Jacob, J. & Baltimore, D. *Nature* **399**, 593–597 (1999).
- Rocha, B. *Nature* **399**, 531–532 (1999).
- Millán, J. L. & Fishman, W. H. *Crit. Rev. Clin. Lab. Sci.* **32**, 1–39 (1995).
- Terao, M. & Mintz, B. *Proc. Natl Acad. Sci. USA* **84**, 7051–7055 (1987).

Maris et al. reply — Mantovani suggests that expression of human placental alkaline phosphatase in activated T cells might stimulate an immune response against PLAP and that the anti-PLAP antibodies generated could lead to the elimination of CD8⁺ T cells expressing large amounts of PLAP on their surface. We therefore tested whether doubly transgenic (GranzymeB-Cre × CD2-STOP-PLAP) mice could mount an antibody response against PLAP during the course of a response against LCMV, and whether the serum of these mice contained antibodies against human PLAP.

We infected doubly transgenic mice ($n=15$) with 5×10^4 plaque-forming units of the Armstrong strain of LCMV and nine days later collected their serum to test for the presence of antibodies against human PLAP in an enzyme-linked immunosorbent

assay (ELISA). None of the fifteen LCMV-infected doubly transgenic mice produced anti-PLAP antibodies, as shown by a representative ELISA from five LCMV-infected (day 9 post-infection) doubly transgenic mice (Fig. 1a).

To find out whether these mice mount a strong antibody response against LCMV, we tested the serum from the same 15 mice for antibodies against LCMV and found that 14 of the 15 mice mounted strong antibody responses against this virus. Figure 1b shows the anti-LCMV antibody levels in the same five mice that were used to obtain the results shown in Fig. 1a — four out of five mice mounted strong anti-LCMV antibody responses.

These results indicate that the low frequency of PLAP-positive CD8⁺ T cells (10%)¹ compared with the LCMV peptide/MHC tetramer-positive CD8⁺ T cells (70%)² observed in the acute phase of the LCMV response is not due to an antibody response against the human PLAP reporter and antibody-mediated elimination of PLAP-expressing CD8⁺ T cells. But we still need to investigate why PLAP marks only a subset of the activated CD8⁺ T cells.

Charles H. Maris*, Joshy Jacob*, David Baltimore†

*Emory University, Atlanta, Georgia 30322, USA

†California Institute of Technology, Pasadena, California 91125, USA

- Jacob, J. & Baltimore, D. *Nature* **399**, 593–597 (1999).
- Murali-Krishna, K. et al. *Immunity* **8**, 177–187 (1998).
- Kuvstak, E. *Enzyme Immunodiagnosis* (Academic, San Diego, 1986).

Microtubule motors in mitosis

David J. Sharp, Gregory C. Rogers & Jonathan M. Scholey

Section of Molecular and Cellular Biology, University of California-Davis, 1 Shields Ave, Davis, California 95616, USA

The mitotic spindle uses microtubule-based motor proteins to assemble itself and to segregate sister chromatids. It is becoming clear that motors invoke several distinct mechanisms to generate the forces that drive mitosis. Moreover, in carrying out its function, the spindle appears to pass through a series of transient steady-state structures, each established by a delicate balance of forces generated by multiple complementary and antagonistic motors. Transitions from one steady state to the next can occur when a change in the activity of a subset of mitotic motors tips the balance.

The mitotic spindle is a bipolar, self-organizing microtubule (MT)-based machine that uses energy liberated from nucleotide hydrolysis to segregate sister chromatids accurately into the daughter cells during cell division^{1,2}. In this review we focus on two inter-related aspects of spindle action: (1) spindle morphogenesis, characterized by the separation of the spindle poles that occurs during spindle assembly, maintenance and elongation, and (2) chromosome movements that occur during the capture of chromosomes, their alignment on the spindle equator and the segregation of sister chromatids (Fig. 1).

Mitotic-spindle function and design

The mechanism of spindle action has long fascinated biologists and during the past 50 years, two critical advances were made. First, the observation that the spindle is a very dynamic structure, superficially resembling a liquid crystal³, led to the discovery that spindle MTs display dynamic instability, undergoing transitions between episodes of polymerization and depolymerization that are coupled to GTP hydrolysis⁴. The dynamic properties of MTs are obviously important for spindle morphogenesis as well as force generation within the spindle⁵. Second, the proposal that mitotic movements could be explained by a 'sliding filament' mechanism⁶ led to the elucidation of the structural organization of spindle MTs^{2,7} (Fig. 2) and a search for mitotic motors that could use ATP hydrolysis to drive the sliding of MTs relative to adjacent MTs or to other spindle structures. These motors form the focus of this review.

MT-based motor proteins form two families of ATP-dependent force-generating enzymes, the kinesins and the dyneins^{8,9}. Following the seminal discovery of two kinesin motors with mitotic functions^{10,11} it has become established that a strikingly large number of MT-based motors participate in spindle action (Table 1). During the past decade, great strides have been made in our understanding of the basic motor mechanisms of mitosis. Surprisingly, it appears that mitotic motors use at least three distinct mechanisms: (1) cross-bridging and sliding MTs relative to adjacent MTs or other structures; (2) transporting specific mitotic cargoes along the surface lattice of spindle MTs; and (3) regulating MT assembly dynamics and coupling movement to MT growth and shrinkage (see Box 1). In addition, analyses of the functional inter-relationships between multiple mitotic motors have revealed that specific mitotic movements are not driven by individual motors but, instead, result from shifts in a dynamic balance of complementary and antagonistic forces generated by multiple motors functioning co-operatively.

Mitotic motor mechanisms

MT sliding. It has long been proposed that force is generated in the spindle by motor-driven MT sliding⁶. For example, motors could crosslink and slide overlapping antiparallel MTs within interpolar MT bundles to position spindle poles in relation to one another. A growing body of evidence indicates that motors actually use two dis-

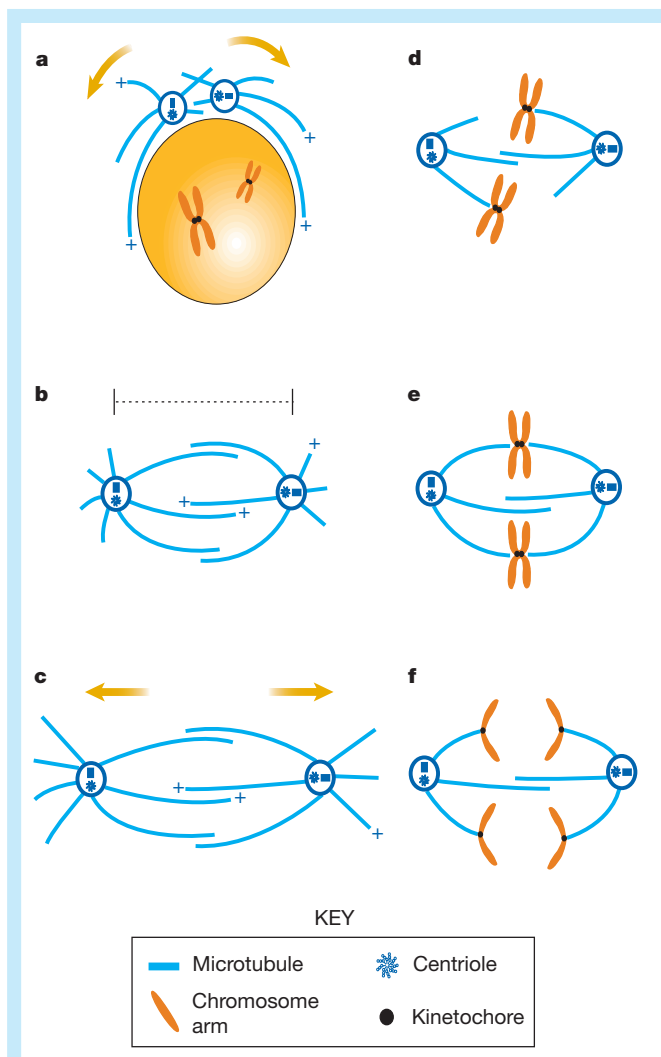


Figure 1 Spindle pole and chromosome movements during mitosis. The activities of MTs and their associated motors are responsible for the carefully orchestrated sequence of movements that underlie mitosis. Early in mitosis (prophase and prometaphase) the bipolar spindle assembles (**a**) and condensed chromosomes are captured by spindle MTs and congress at the spindle equator (**d**). By metaphase, the spindle consists of two partially interdigitating arrays of MTs emanating from the duplicated centrosomes that form the two spindle poles (**b**). Chromosomes lie on the equator with pairs of identical sister chromatids facing opposite spindle poles (**e**). Chromatid segregation involves both the movement of sister chromatids to opposite spindle poles (**f**; anaphase A) and the elongation of the spindle itself (**c**; anaphase B). Mechanistically it is convenient to think of these events in terms of spindle assembly, maintenance and elongation (left), and chromosome capture, congression and segregation (right). MT dynamics and length adjustments are in refs 2, 5.

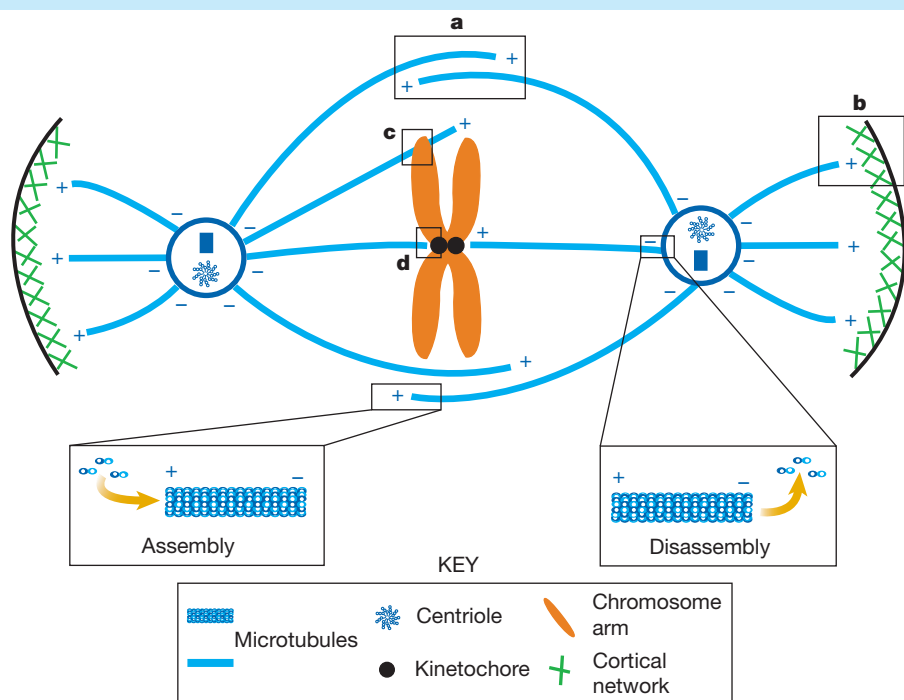


Figure 2 Structure and polarity patterns of MTs in the metaphase spindle. Structural studies have revealed that spindle MTs are all orientated with their fast-growing (plus) ends distal to the spindle poles, but are organized into four functionally distinct subsets. Interpolar MT bundles exert forces capable of moving opposite spindle poles relative to one another. Kinetochore MT bundles move chromosomes relative to spindle poles. Astral MTs link spindle poles to the cell cortex and contribute to the separation of spindle poles and the positioning of the spindle relative to the cell cortex. A fourth set of MTs link centrosomes to chromosome arms. It is now clear that the spindle contains mitotic motors that use ATP hydrolysis to generate forces relative to all four sets of spindle MTs. **a–d**, Putative sites of mitotic-motor activity shown in detail in Box 1.

tinct forms of this ‘sliding filament mechanism’ during mitosis: sliding adjacent MTs in relation to one another and sliding astral MTs in relation to a stationary cell cortex (Box 1).

The strongest evidence that mitotic motors function by MT–MT sliding is from analyses of the fast-growing (plus)-end-directed bipolar (bimC) kinesins (Fig. 3). Members of this kinesin subfamily form bipolar homotetramers, with two MT-motor domains positioned at opposite ends of a central rod^{12–14}, indicating that they generate force along adjacent MTs simultaneously. Moreover, the *Drosophila* bipolar kinesin crosslinks adjacent spindle MTs within interpolar MT bundles *in situ*¹⁵, and functional inhibition of bipolar kinesins results, uniformly across phyla, in the formation of spindles with abnormally close poles, *in vivo*^{10,16–22}. The bipolar kinesins are also structurally similar to class II myosins, which crosslink and slide antiparallel actin filaments to contract the muscle sarcomere²³. These data strongly indicate that these mitotic motors slide antiparallel spindle MTs apart to generate ‘outward’ forces between spindle poles.

It appears that another class of mitotic MT–MT sliding motors function analogously to the axonemal forms of dynein that drive ciliary motility. These motors include members of the carboxy-terminal and CHO1/MKLP1 kinesins as well as dynein/dynactin. Like axonemal dynein, they contain nucleotide-insensitive MT-binding sites distal to their motor domains^{11,24–30}, allowing them to bind an MT as cargo and slide it in relation to a neighbouring MT. This hypothesis is supported by the observation that members of all three classes of motors can crosslink MTs into bundles, *in vitro* or in extracts^{29,31–33}. Moreover, there is evidence that the C-terminal and CHO1/MKLP1 kinesins position spindle poles and organize interpolar MT bundles, *in vivo*, consistent with an antiparallel-MT bundling and sliding mechanism^{34–36}.

MT-sliding forces within the spindle appear not to be confined to MT–MT sliding. It is probable that mitotic motors exert force by crosslinking and sliding astral MTs in relation to a dense cortex of actin filaments and associated proteins that forms just beneath the

Table 1 Sites of action and functional directionality of mitotic motors

Spindle position	Motor family	Functional directionality	Mechanism of action
Midzone/interpolar MTs (Box 1a)	Bipolar kinesin	Pole–pole: outward	MT–MT sliding/bundling
	C-terminal kinesin	Pole–pole: inward	
	CHO1/MKLP1	Pole–pole: outward/tight bundling	
Cortical (Box 1b)	Dynein/dynactin	Pole–pole: outward	Actin–MT sliding
	Dynein/dynactin	Chromosome: poleward	Kinetochore transport
		Pole–pole: inward	
Kinetochore (Box 1d)	MCAK/XKCM1	Chromosome: poleward	MT disassembly
		Pole–pole: inward	
	CENP-E	Chromosome: poleward/plateward	Kinetochore transport/MT plus-end anchor
		Pole–pole: inward/outward	
Chromosome (Box 1c)	Chromokinesins	Chromosome: plateward	Chromosome transport
		Pole–pole: outward	

cell surface. Dynein/dynactin, for example, which positions mitotic spindle poles³⁷, clearly localizes to the cell cortex during mitosis in many systems^{34,38–40} and thus is positioned appropriately to capture and slide MTs extending away from the central spindle (Fig. 4). Moreover, although dynein-mediated actin–MT sliding has not been directly demonstrated, a component of the dynactin complex does bind the actin-binding protein, spectrin, *in vitro*⁴¹.

In addition to the role of MT-sliding motors in positioning spindle poles, motor-driven MT bundling and sliding may also focus MTs at spindle poles⁴². For example, the *Drosophila* C-terminal kinesin, Ncd, can 'zip' together MTs at their minus ends in the anastral (centrosome-free) spindles that form during female meiosis⁴³. Also dynein/dynactin is involved in the formation of both anastral spin-

dles and MT asters that arise in cell-free M-phase extracts under certain experimental conditions^{33,44}. Though it is possible that these motors also organize mitotic spindle poles, there is no clear experimental data from living mitotic cells supporting this hypothesis. This may be because centrosomes anchor MTs at the poles of mitotic spindles, making this activity less obvious.

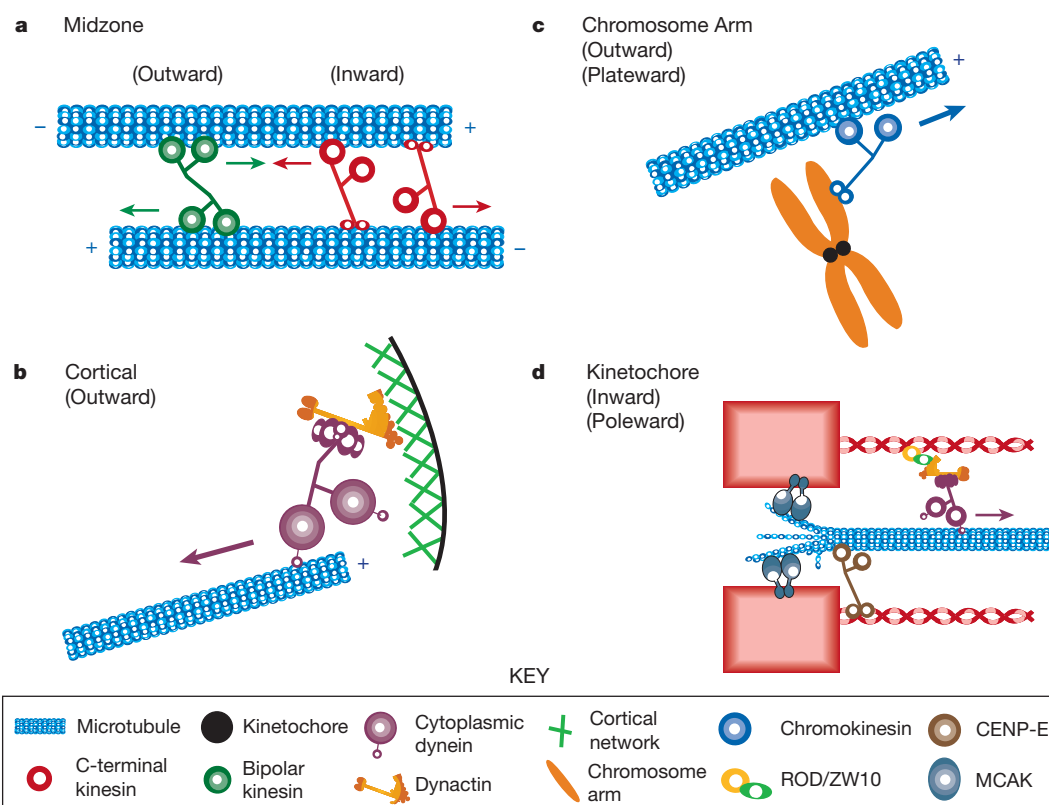
Transport of chromosomes along spindle MTs. The organization of spindle MTs has implications for how motors position chromosomes; plus-end-directed motors could drive chromosomes toward the metaphase plate while minus-end-directed motors could drive them toward the spindle poles. The last decade has seen the identification of multiple mitotic motors that are potentially involved in chromosome movements (Table 1). Interestingly, many of these

Box 1

Sites and mechanisms of action for mitotic motors

Given the structural organization of the spindle and the polarity patterns of spindle MTs, it is probable that multiple motors with distinct localizations, transport properties and mechanisms provide the cell with a precisely controlled, co-ordinated means for driving coherent mitotic movements. The specific sites in the spindle where motors are likely to act and their mechanisms are shown. **a**, Midzone. Motors can crosslink and slide antiparallel MTs in the spindle midzone using one of two mechanisms: (1) motors may form bipolar structures with motor domains positioned at both ends (bipolar kinesins) or (2) asymmetric structures that bind one MT as cargo (through nucleotide-insensitive MT-binding sites) and transport it along neighbouring MTs. Owing to the MT-polarity patterns in the spindle, antiparallel MT sliding would allow plus-end-directed motors to push the poles apart and minus-end-directed motors to pull them together. **b**, Cortex. Motors can crosslink and slide astral MTs in relation to an actin matrix at the cell cortex. This would allow minus-end-directed motors such as cytoplasmic dynein to pull the spindle poles apart and position the entire spindle within the cell. **c**, Chromosome arms.

Motors can bind chromosome arms as cargo and transport them along adjacent spindle MTs. These forces would position chromosomes on the spindle and also position the spindle poles in relation to the chromosomal mass. Plus-end-directed motors would drive chromosomes towards the metaphase plate (plateward) and push the poles away from chromosomes (outward). **d**, Kinetochore. Kinetochore motility involves two distinct motor mechanisms; (1) transport along the MT surface lattice and (2) coupling of movement with MT assembly dynamics. Minus-end-directed motors on the fibrous corona extending away from the chromosomes can transport kinetochores along the MT surface towards the spindle poles (poleward). Motors positioned in the interior of the kinetochore can de-stabilize kinetochore MTs, causing them to shorten towards the poles. Simultaneously, plus-end-directed motors (that transport kinetochores plateward until the plus-end of the MT is encountered) can anchor the kinetochore to the shortening plus-end of the MT to couple poleward kinetochore motility with MT shrinkage.



appear to act similarly to vesicle transport motors, by transporting specific chromosomal regions as cargo along spindle MTs (Box 1).

A major site for chromosome motility is the kinetochore, a multi-protein complex that can move bi-directionally along spindle MTs (Fig. 4)⁴⁶. Consistent with the hypothesis that kinetochores move by motor-driven transport along the surface lattice of MTs, both the minus-end-directed cytoplasmic dynein and the plus-end-directed kinesin, CENP-E, immunolocalize to kinetochores *in situ*^{47–49}. Moreover, functional analyses indicate that both of these motors are important for properly positioning chromosomes on the spindle in a manner consistent with their transport properties; the inhibition of CENP-E results in defects in chromosome congression⁵⁰, and several mutations in genes encoding components of the dynein complex result in defects in chromosome segregation^{51–53}. Unfortunately, because of the fundamental importance of kinetochore motility in mitosis, there is likely to be a good deal of redundancy in the mechanisms driving its movement; which may explain why the inhibition of any single motor cannot completely abolish kinetochore movement.

It is also clear that chromosome positioning is determined by plateward or 'polar ejection forces' generated directly on chromosome arms⁵⁴. Although the mechanism for this is not understood, analyses of several mitotic motors indicate they have the capacity to do this by a mechanism similar to that of the kinetochore motors described above. A number of presumptively plus-end-directed mitotic motors, including members of the chromokinesin subfamily⁵⁵, KLP38B in *Drosophila*^{56–58} and kid from *Homo sapiens*⁵⁹, contain specific DNA- or chromatin-binding motifs and concentrate on the non-kinetochore regions of chromosomes. Thus, they can bind chromosome arms as cargo and transport them toward the metaphase plate, generating 'outward' forces on spindle poles. Recent functional analyses of specific chromosomal motors support this hypothesis^{56–58}.

Regulation of MT assembly dynamics. Besides transporting kinetochores along spindle MTs, mitotic motors probably couple MT assembly dynamics to kinetochore motility. The best evidence supporting such a function is from analyses of the movement of chromosome fragments along MTs nucleated by lysed *Tetrahymena* cells. In these studies, the presence of anti-CENP-E antibodies significantly inhibited the normal movement of chromosomes towards MT minus-ends during the induction of MT depolymerization⁶⁰. These antibodies had no overt effects on the assembly dynamics of the MTs themselves, but reduced the ability of chromosomes to associate with the ends of shortening MTs. These surprising results indicate that CENP-E could use its plus-end-directed transport properties in two

ways; to transport kinetochores toward the MT plus-ends through metaphase and subsequently to anchor kinetochores to the shortening plus-ends of MTs during anaphase (Box 1).

More recent data support the notion that motors can also regulate MT disassembly directly. Two related *Xenopus laevis* motors containing motor domains in the interior of their polypeptides, XKCM1 and XKIF2, destabilize MTs at their ends, *in vitro*^{61,62}. Furthermore, XKCM1 and its mammalian homologue, MCAK, localize to kinetochores during mitosis, and functional analyses indicate roles in both spindle assembly and chromosome movements^{61,64}. Thus, MCAK/XKCM1 motors could induce the disassembly of kinetochore MTs which, in concert with the plus-end-anchoring activity of CENP-E, could transduce MT shortening into poleward forces on chromosomes.

It is, perhaps, not surprising that cells use multiple motors and mechanisms for chromosome motility. The orchestration of these mechanisms ensures that the force applied to chromosomes by the spindle is highly precise and can be regulated at various levels. Also, by coupling chromosome movements to multiple mechanisms, cells have evolved an efficient means of exploiting the intrinsic polarity properties of MTs with a measure of redundancy that may allow slight obstacles to be overcome.

Functional relationships between mitotic motors

An important issue facing mitosis researchers is why the mitotic spindle contains so many motors, some with rather similar properties. The yeast, *Saccharomyces cerevisiae*, for example, contains two members of the bipolar kinesin family, Cin8p and Kip1p. Although cells containing double knockouts of these motors display defects in the relative positioning of spindle poles, cells with single knockouts have no observable phenotype, leading to the proposal that similar mitotic motors have functional redundancy^{17,18}. However, recent live-cell quantitative analyses reveal subtle but significant differences in when these motors exert their influence on spindle pole separation⁶⁵. Thus, these similar motors perform complementary but not entirely overlapping functions. Also, bipolar and C-terminal kinesins function antagonistically^{22,26,66–68}. These data indicate that cells use multiple mitotic motors in parallel to generate a delicate balance of complementary and antagonistic forces.

Transitions between steady-state spindle structures. These concepts, derived largely from yeast genetics¹, have recently been applied to the pathway of spindle pole separation in the *Drosophila* early embryo where it is possible to visualize and quantify spindle

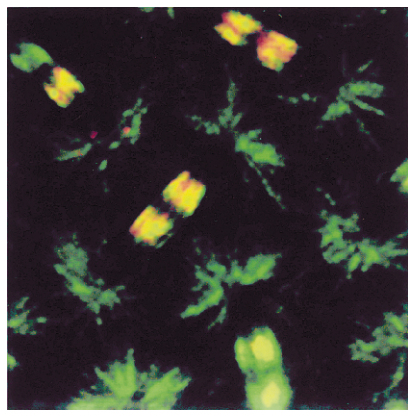


Figure 3 Immunofluorescence micrograph showing the bipolar kinesin, in the midzone of *Drosophila* embryonic spindles. The kinesin motor is shown in red, spindle microtubules are shown in green and overlap is yellow.

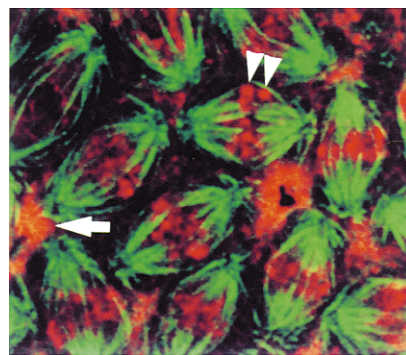


Figure 4 Immunofluorescence micrograph showing the localization of the dynein motor protein to kinetochores (double arrowheads) and cortical actin networks (arrow) in *Drosophila* embryos.

Box 2

Motor pathway for spindle assembly, maintenance and elongation

Numerous studies in a variety of systems have indicated that specific movements of the mitotic spindle are not driven by individual mitotic motors acting on their own. Indeed, it is now known that multiple complementary and antagonistic motors function simultaneously within the spindle and recent analyses of the relative movements of spindle poles in *Drosophila* in the presence and absence of specific motor inhibitors have provided insights into how multiple motors work together.

The right panels show plots of spindle pole separation as a function of time from control *Drosophila* early embryos. Note that the rate of spindle pole separation is not linear but occurs in bursts interspersed with quiescent pauses. Because these pauses are highly stereotypical and occur at characteristic stages of mitosis and spindle lengths, we propose that they represent a series of transient steady-state structures, which result when all active motors within the spindle precisely balance one another. The transition between steady states, which would occur when the activity of a subset of motors changes, could be the driving force behind specific mitotic movements. While on the surface such a mechanism may appear inefficient, the use of multiple motors working in parallel provides a very precise means of controlling the strength and directionality of the force applied to spindle poles throughout mitosis.

Parallel analyses of the rates of spindle pole separation in the presence of specific inhibitors against bipolar kinesins, C-terminal kinesins, and dynein/dynactin applied in various combinations (assessed in a context provided by studies of these motors in other systems) have led to the motor pathway for spindle assembly, maintenance and elongation shown in the left panels. In this pathway, nearly all spindle pole movements can be accounted for by the dynamic interplay between these MT-sliding motors.

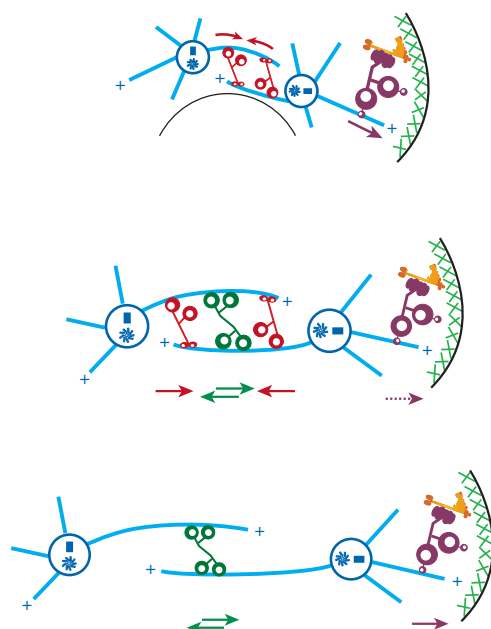
Interphase–prophase. The initial assembly of the spindle during interphase and prophase involves a dynamic balance between dynein on the cortex and C-terminal kinesins in the spindle midzone, pulling the poles apart and together, respectively. (Through this stage in mitosis, the bipolar kinesin is sequestered within the nucleus and does not appear to contribute to spindle pole positioning.) At the outset of spindle assembly, the extent of antiparallel MT overlap in the

midzone is minimal, limiting the relative strength of the forces generated by C-terminal kinesins. However, as the poles separate further and growing MTs continue to interdigitate, the inward force generated by the C-terminal kinesin gradually increases until it balances the outward force generated by dynein, resulting in a steady-state structure (SS1). This balance is tipped just before nuclear envelope breakdown.

Prometaphase–metaphase. Another steady state is achieved immediately following nuclear envelope breakdown (SS2). At this point astral MTs are minimal. This steady state may result from interactions between bipolar kinesins, C-terminal kinesins and interpolar MT bundles²². Within a minute of nuclear envelope breakdown, however, the forces within the spindle shift again. This is almost certainly due to cortical dynein plus bipolar kinesin activity, overwhelming the inward forces generated by C-terminal kinesins and tipping the balance of forces in the outward direction. The result is the elongation of the spindle and establishment of the metaphase spindle steady-state structure (SS3).

Anaphase B. Tension generated by antagonistic 'inward' and 'outward' forces within the metaphase spindle keeps it coiled like a spring. This spring is released at the onset of anaphase, causing a final elongation of the spindle. Interestingly, the trigger for this movement is not the activation of a third outward motor, but, instead, the inactivation of the inward forces generated by the C-terminal kinesin. Just before the disassembly of the spindle, a fourth steady-state structure forms at telophase (SS4).

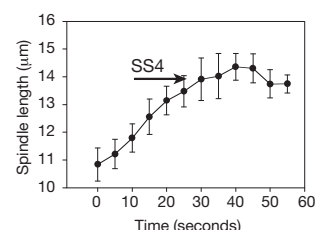
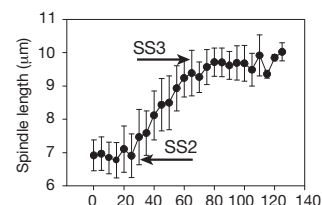
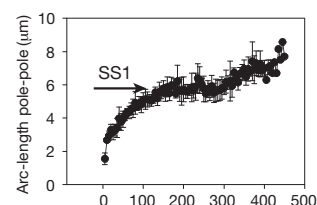
However, this is clearly an incomplete description of spindle morphogenesis. Mitosis in higher eukaryotes is known to involve more than three motors, many of which are essential on their own, and the dynamic properties of MTs themselves. We propose that these also contribute to the dynamic balance of forces within the spindle. Spindle morphology and the rate of spindle pole separation at any point during mitosis represents the sum of all active motors; or the difference between the overall inward versus outward forces generated on the spindle poles. A time-series movie showing mitosis in *Drosophila* embryos is presented in the Supplementary Information.



Interphase–Prophase

Prometaphase–Metaphase

Anaphase B



pole movements in the presence and absence of specific motor inhibitors³⁴. The results argue that mitotic movements are not driven by individual motors acting alone. Instead, the spindle appears to progress through a series of transient steady-state structures as multiple complementary and antagonistic motors precisely balance one another. Specific mitotic movements are therefore driven by changes in the relative strength of subsets of motors which induce the transition from one steady state to the next (Box 2).

Transient motor-generated steady-state structures may also serve as a mechanism for chromosome positioning during chromosome capture, congression and segregation. Although this has not been demonstrated experimentally, progression through this pathway could plausibly be signalled by the density and polarity patterns of MTs surrounding chromosome arms and kinetochores. The response of a chromosome to its position would, in turn, be determined by the relative strength of the poleward versus plateward forces generated by motors positioned on these structures. Thus, chromosomes would always tend to move towards a specific steady-state or balance position, which could be altered at specific stages of the cell cycle by alterations in the activity of specific motors as well as subtle changes in spindle and chromosome structure.

Cell-cycle regulation of spindle steady states. There is a good deal of evidence supporting a dynamic balance of complementary and antagonistic motor-generated forces during mitotic spindle formation and function, but how this balance is linked to the cell cycle is not known. It seems most probable that this is controlled, either directly or indirectly, by the oscillatory activity of cyclin-dependent kinases (cdks) and other regulatory kinases and phosphatases. Support for this hypothesis comes from studies showing that the expression of non-degradable cyclin constructs in *Drosophila* embryos 'freezes' spindles in specific steady-state configurations⁶⁹. Moreover, several mitotic motors serve as substrates for different cell-cycle regulators^{21,70–73}. Changes in the phosphorylation state of mitotic motors could alter their localization, transport properties, structural configuration or expression levels, in each case influencing the net force within the spindle, and potentially driving it from one steady state to the next.

Future perspectives

This review highlights recent results illuminating the mechanisms by which individual mitotic motors act and also provides insights into how the activities of multiple mitotic motors are integrated into the pathways of spindle pole positioning and chromosome movements. Our major thesis is that steady-state spindle structures depend on a balance of forces generated by multiple complementary and antagonistic motors, and tipping this balance by up- or down-regulating subsets of motors drives specific mitotic movements. Though the models presented are plausible, much work needs to be done to test, refine and correct them. Many questions remain concerning individual mitotic motors. For example, the predicted ability of purified bipolar kinesins to crosslink and slide antiparallel MTs apart *in vitro* has not been demonstrated, and the significance of the differences in structural organization between the bipolar and C-terminal kinesins in driving MT–MT sliding is unclear. Furthermore, little is known of how individual mitotic motors are targeted to their site of action in the spindle. The models describing the functional co-ordination of multiple mitotic motors rely on methods (for example, mutants or antibody micro-injection) for rapidly inactivating specific motor function and observing the consequences for spindle action. Arguably the best reagents available to accomplish this are fast-acting conditional loss-of-function mutants, but few studies use reagents of this level of efficacy, and other approaches are needed. A promising alternative might be the use of chemical genetics to identify cell-penetrating chemical inhibitors that can rapidly and reversibly inactivate mitotic motors in cells⁷⁴. Further work may produce quantitative models that explain how multiple motors (plus MT dynamics) generate forces of the magnitude seen in spindles, and how these forces

influence spindle structure⁷⁵. The models of multiple motor cooperation in the positioning of spindle poles and chromosomes raise the question of how the activities of specific motors or sets of motors are increased or decreased at specific times during mitosis. In this context, it will be important to learn how mitotic motors respond to chemical (for example, phosphorylation–dephosphorylation) and mechanical (for example, tension exerted by changes in the activity of complementary or antagonistic motors) signals, as well as how they are regulated in the broad context of the cell cycle. Finally, our new understanding of the mechanisms of mitotic motor action in the spindle might provide insights into some of the dysfunctions in spindle behaviour and chromosome segregation that characterize certain birth defects and cancer.

Note added in proof. Very recent advances in mitotic motor research were reviewed in ref. 76. Further support for the hypothesis that chromosome positioning is determined by polar ejection forces generated on chromosome arms, discussed in the section titled 'Transport of chromosomes along spindle MTs' was recently published^{77,78}. □

- Hoyt, M. A. Geiser, J. R. Genetic analysis of the mitotic spindle. *Annu. Rev. Genet.* **30**, 7–33 (1996).
- McIntosh, J. R. & McDonald, K. L. The mitotic spindle. *Sci. Am.* **261**, 48–56 (1989).
- Inoue, S. & Sato, H. Cell motility by labile association of molecules. The nature of mitotic spindle fibers and their role in chromosome movement. *J. Gen. Physiol.* **50** (Suppl.), 259–292 (1967).
- Mitchison, T. & Kirschner, M. Dynamic instability of microtubule growth. *Nature* **312**, 237–242 (1984).
- Inoue, S. & Salmon, E. D. Force generation by microtubule assembly/disassembly in mitosis and related movements. *Mol. Biol. Cell* **6**, 1619–1640 (1995).
- McIntosh, J. R., Hepler, P. K. & Van Wie, D. G. Model for mitosis. *Nature* **224**, 659–663 (1969).
- McDonald, K. L., Edwards, M. K. & McIntosh, J. R. Cross-sectional structure of the central mitotic spindle of *Diatoma vulgare*. Evidence for specific interactions between antiparallel microtubules. *J. Cell Biol.* **83**, 443–461 (1979).
- Vale, R. D. & Fletterick, R. J. The design plan of kinesin motors. *Annu. Rev. Cell Dev. Biol.* **13**, 745–777 (1997).
- Holzbaur, E. L. & Vallee, R. B. DYNEINS: molecular structure and cellular function. *Annu. Rev. Cell Biol.* **10**, 339–372 (1994).
- Enos, A. P. & Morris, N. R. Mutation of a gene that encodes a kinesin-like protein blocks nuclear division in *A. nidulans*. *Cell* **60**, 1019–1027 (1990).
- Meluh, P. B. & Rose, M. D. KAR3, a kinesin-related gene required for yeast nuclear fusion. *Cell* **60**, 1029–1041 (1990); erratum *ibid* **61**, 548.
- Cole, D. G., Saxton, W. M., Sheehan, K. B. & Scholey, J. M. A "slow" homotetrameric kinesin-related motor protein purified from *Drosophila* embryos. *J. Biol. Chem.* **269**, 22913–22916 (1994).
- Kashina, A. S. *et al.* A bipolar kinesin. *Nature* **379**, 270–272 (1996).
- Gordon, D. M. & Roof, D. M. The kinesin-related protein Kip1p of *Saccharomyces cerevisiae* is bipolar. *J. Biol. Chem.* **274**, 28779–28786 (1999).
- Sharp, D. J. *et al.* The bipolar kinesin, KLP61F, cross-links microtubules within interpolar microtubule bundles of *Drosophila* embryonic mitotic spindles. *J. Cell Biol.* **144**, 125–138 (1999).
- Hagan, I. & Yanagida, M. Novel potential mitotic motor protein encoded by the fission yeast cut7+ gene. *Nature* **347**, 563–566 (1990).
- Hoyt, M. A., He, L., Loo, K. K. & Saunders, W. S. Two *Saccharomyces cerevisiae* kinesin-related gene products required for mitotic spindle assembly. *J. Cell Biol.* **118**, 109–120 (1992).
- Roof, D. M., Meluh, P. B. & Rose, M. D. Kinesin-related proteins required for assembly of the mitotic spindle. *J. Cell Biol.* **118**, 95–108 (1992).
- Sawin, K. E., LeGuellec, K., Philippe, M. & Mitchison, T. J. Mitotic spindle organization by a plus-end-directed microtubule motor. *Nature* **359**, 540–543 (1992).
- Heck, M. M. *et al.* The kinesin-like protein KLP61F is essential for mitosis in *Drosophila*. *J. Cell Biol.* **123**, 665–679 (1993).
- Blangy, A. *et al.* Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation *in vivo*. *Cell* **83**, 1159–1169 (1995).
- Sharp, D. J., Yu, K. R., Sisson, J. C., Sullivan, W. & Scholey, J. M. Antagonistic microtubule-sliding motors position mitotic centrosomes in *Drosophila* early embryos. *Nature Cell Biol.* **1**, 51–54 (1999).
- Huxley, H. E. Sliding filaments and molecular motile systems. *J. Biol. Chem.* **265**, 8347–8350 (1990).
- Chandra, R., Salmon, E. D., Erickson, H. P., Lockhart, A. & Endow, S. A. Structural and functional domains of the *Drosophila* ncd microtubule motor protein. *J. Biol. Chem.* **268**, 9005–9013 (1993).
- Kuriyama, R. *et al.* Characterization of a minus end-directed kinesin-like motor protein from cultured mammalian cells. *J. Cell Biol.* **129**, 1049–1059 (1995).
- Pidoux, A. L., LeDizet, M. & Cande, W. Z. Fission yeast pkl1 is a kinesin-related protein involved in mitotic spindle function. *Mol. Biol. Cell* **7**, 1639–1655 (1996).
- Karabay, A. & Walker, R. A. Identification of microtubule binding sites in the Ncd tail domain. *Biochemistry* **38**, 1838–1849 (1999).
- Narasimulu, S. B. & Reddy, A. S. Characterization of microtubule binding domains in the *Arabidopsis* kinesin-like calmodulin binding protein. *Plant Cell* **10**, 957–965 (1998).
- Kuriyama, R. *et al.* Heterogeneity and microtubule interaction of the CHO1 antigen, a mitosis-specific kinesin-like protein. Analysis of subdomains expressed in insect sf9 cells. *J. Cell Sci.* **107**, 3485–3499 (1994).
- Waterman-Storer, C. M., Karki, S. & Holzbaur, E. L. The p150Glued component of the dynactin complex binds to both microtubules and the actin-related protein centractin (Arp-1). *Proc. Natl Acad. Sci. USA* **92**, 1634–1638 (1995).
- McDonald, H. B., Stewart, R. J. & Goldstein, L. S. The kinesin-like ncd protein of *Drosophila* is a minus end-directed microtubule motor. *Cell* **63**, 1159–1165 (1990).

32. Nislow, C., Lombillo, V. A., Kuriyama, R. & McIntosh, J. R. A plus-end-directed motor enzyme that moves antiparallel microtubules in vitro localizes to the interzone of mitotic spindles. *Nature* **359**, 543–547 (1992).
33. Verde, F., Berrez, J. M., Antony, C. & Karsenti, E. Taxol-induced microtubule asters in mitotic extracts of *Xenopus* eggs: requirement for phosphorylated factors and cytoplasmic dynein. *J. Cell Biol.* **112**, 1177–1187 (1991).
34. Sharp, D. J. *et al.* Functional coordination of three mitotic motors in *Drosophila* embryos. *Mol. Biol. Cell* **11**, 241–253 (2000).
35. Adams, R. R., Tavares, A. A., Salzberg, A., Bellen, H. J. & Glover, D. M. *pavarotti* encodes a kinesin-like protein required to organize the central spindle and contractile ring for cytokinesis. *Genes Dev.* **12**, 1483–1494 (1998).
36. Raich, W. B., Moran, A. N., Rothman, J. H. & Hardin, J. Cytokinesis and midzone microtubule organization in *Caenorhabditis elegans* require the kinesin-like protein ZEN-4. *Mol. Biol. Cell* **9**, 2037–2049 (1998).
37. Sharp, D. J., Rogers, G. C. & Scholey, J. M. Roles of motor proteins in building microtubule-based structures: a basic principle of cellular design. *Biochim. Biophys. Acta* **1496**, 128–141 (2000).
38. Yeh, E., Skibbens, R. V., Cheng, J. W., Salmon, E. D. & Bloom, K. Spindle dynamics and cell cycle regulation of dynein in the budding yeast, *Saccharomyces cerevisiae*. *J. Cell Biol.* **130**, 687–700 (1995).
39. Busson, S., Dujardin, D., Moreau, A., Dompierre, J. & De Mey, J. R. Dynein and dynactin are localized to astral microtubules and at cortical sites in mitotic epithelial cells. *Curr. Biol.* **8**, 541–544 (1998).
40. O'Connell, C. B. & Wang, Y. Mammalian spindle orientation and position respond to changes in cell shape in a dynein-dependent fashion. *Mol. Biol. Cell* **11**, 1765–1764 (2000).
41. Holleran, E. A., Tokito, M. K., Karki, S. & Holzbaur, E. L. Centractin (ARP1) associates with spectrin revealing a potential mechanism to link dynactin to intracellular organelles. *J. Cell Biol.* **135**, 1815–1829 (1996).
42. Merdes, A. & Cleveland, D. W. Pathways of spindle pole formation: different mechanisms; conserved components. *J. Cell Biol.* **138**, 953–956 (1997).
43. Matthies, H. J., McDonald, H. B., Goldstein, L. S. & Theurkauf, W. E. Anastral meiotic spindle morphogenesis: role of the non-claret disjunctional kinesin-like protein. *J. Cell Biol.* **134**, 455–464 (1996).
44. Heald, R. *et al.* Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* **382**, 420–425 (1996).
45. Rieder, C. L. & Salmon, E. D. The vertebrate cell kinetochore and its roles during mitosis. *Trends Cell Biol.* **8**, 310–8 (1998).
46. Hyman, A. A. & Mitchison, T. J. Two different microtubule-based motor activities with opposite polarities in kinetochores. *Nature* **351**, 206–211 (1991).
47. Yen, T. J., Li, G., Schaar, B. T., Szilak, I. & Cleveland, D. W. CENP-E is a putative kinetochore motor that accumulates just before mitosis. *Nature* **359**, 536–539 (1992).
48. Steuer, E. R., Wordeman, L., Schroer, T. A. & Sheetz, M. P. Localization of cytoplasmic dynein to mitotic spindles and kinetochores. *Nature* **345**, 266–268 (1990).
49. Pfarr, C. M. *et al.* Cytoplasmic dynein is localized to kinetochores during mitosis. *Nature* **345**, 263–265 (1990).
50. Wood, K. W., Sakowicz, R., Goldstein, L. S. & Cleveland, D. W. CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome alignment. *Cell* **91**, 357–366 (1997).
51. Starr, D. A., Williams, B. C., Hays, T. S. & Goldberg, M. L. ZW10 helps recruit dynactin and dynein to the kinetochore. *J. Cell Biol.* **142**, 763–774 (1998).
52. Bowman, A. B. *et al.* *Drosophila* roadblock and *Chlamydomonas* LC7: a conserved family of dynein-associated proteins involved in axonal transport, flagellar motility, and mitosis. *J. Cell Biol.* **146**, 165–180 (1999).
53. Lee, S., Wisniewski, J. C., Dentler, W. L. & Asai, D. J. Gene knockouts reveal separate functions for two cytoplasmic dyneins in *Tetrahymena thermophila*. *Mol. Biol. Cell* **10**, 771–784 (1999).
54. Rieder, C. L. & Salmon, E. D. Motile kinetochores and polar ejection forces dictate chromosome position on the vertebrate mitotic spindle. *J. Cell Biol.* **124**, 223–233 (1994).
55. Wang, S. Z. & Adler, R. Chromokinesin: a DNA-binding, kinesin-like nuclear protein. *J. Cell Biol.* **128**, 761–768 (1995).
56. Molina, I. *et al.* A chromatin-associated kinesin-related protein required for normal mitotic chromosome segregation in *Drosophila*. *J. Cell Biol.* **139**, 1361–1371 (1997).
57. Ruden, D. M., Cui, W., Sollars, V. & Alterman, M. A *Drosophila* kinesin-like protein, Klp38B, functions during meiosis, mitosis, and segmentation. *Dev. Biol.* **191**, 284–296 (1997).
58. Alpay, L. *et al.* KLP38B: a mitotic kinesin-related protein that binds PP1. *J. Cell Biol.* **138**, 395–409 (1997).
59. Tokai, N. *et al.* Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and the mitotic spindle. *Embo J* **15**, 457–467 (1996).
60. Lombillo, V. A., Nislow, C., Yen, T. J., Gelfand, V. I. & McIntosh, J. R. Antibodies to the kinesin motor domain and CENP-E inhibit microtubule depolymerization-dependent motion of chromosomes *in vitro*. *J. Cell Biol.* **128**, 107–115 (1995).
61. Walczak, C. E., Mitchison, T. J. & Desai, A. XKCM1: a *Xenopus* kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. *Cell* **84**, 37–47 (1996).
62. Desai, A., Verma, S., Mitchison, T. J. & Walczak, C. E. Kin I kinesins are microtubule-destabilizing enzymes. *Cell* **96**, 69–78 (1999).
63. Wordeman, L. & Mitchison, T. J. Identification and partial characterization of mitotic centromere-associated kinesin, a kinesin-related protein that associates with centromeres during mitosis. *J. Cell Biol.* **128**, 95–104 (1995).
64. Maney, T., Hunter, A. W., Wagenbach, M. & Wordeman, L. Mitotic centromere-associated kinesin is important for anaphase chromosome segregation. *J. Cell Biol.* **142**, 787–801 (1998).
65. Straight, A. F., Sedat, J. W. & Murray, A. W. Time-lapse microscopy reveals unique roles for kinesins during anaphase in budding yeast. *J. Cell Biol.* **143**, 687–94 (1998).
66. Saunders, W. S. & Hoyt, M. A. Kinesin-related proteins required for structural integrity of the mitotic spindle. *Cell* **70**, 451–458 (1992).
67. O'Connell, M. J., Meluh, P. B., Rose, M. D. & Morris, N. R. Suppression of the bimC4 mitotic spindle defect by deletion of klpA, a gene encoding a KAR3-related kinesin-like protein in *Aspergillus nidulans*. *J. Cell Biol.* **120**, 153–162 (1993).
68. Mountain, V. *et al.* The kinesin-related protein, HSET, opposes the activity of Eg5 and cross-links microtubules in the mammalian mitotic spindle. *J. Cell Biol.* **147**, 351–66 (1999).
69. Follette, P. J. & O'Farrell, P. H. Cdk and the *Drosophila* cell cycle. *Curr. Opin. Genet. Dev.* **7**, 17–22 (1997).
70. Sawin, K. E. & Mitchison, T. J. Mutations in the kinesin-like protein Eg5 disrupting localization to the mitotic spindle. *Proc. Natl Acad. Sci. USA* **92**, 4289–4293 (1995).
71. Lee, K. S., Yuan, Y. L., Kuriyama, R. & Erikson, R. L. Plk is an M-phase-specific protein kinase and interacts with a kinesin-like protein, CHO1/MKLP-1. *Mol. Cell Biol.* **15**, 7143–7151 (1995).
72. Chan, G. K., Jablonski, S. A., Sudakin, V., Hittle, J. C. & Yen, T. J. Human BUBR1 is a mitotic checkpoint kinase that monitors CENP-E functions at kinetochores and binds the cyclosome/APC. *J. Cell Biol.* **146**, 941–54 (1999).
73. Giet, R., Uzbekov, R., Cubizolles, F., Le Guellec, K. & Prigent, C. The *Xenopus laevis* aurora-related protein kinase pEg2 associates with and phosphorylates the kinesin-related protein XIEg5. *J. Biol. Chem.* **274**, 15005–13 (1999).
74. Mayer, T. U. *et al.* Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* **286**, 971–974 (1999).
75. Nicklas, R. B. The forces that move chromosomes in mitosis. *Annu. Rev. Biophys. Chem.* **17**, 431–449 (1988).
76. Heald, R. Motor function in the mitotic spindle. *Cell* **102**, 399–402 (2000).
77. Funabiki, H. & Murray, A. W. The *Xenopus* Chromokinesin, XKid, is essential for metaphase chromosome alignment and must be upgraded to allow anaphase chromosome movement. *Cell* **102**, 411–424 (2000).
78. Antonio, C. *et al.* XKid, a chromokinesin required for chromosome alignment on the metaphase plate. *Cell* **102**, 525–435 (2000).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>).

Correspondence should be addressed to J.M.S. (e-mail: jmscholey@ucdavis.edu).

Nicastrin modulates presenilin-mediated *notch*/*glp-1* signal transduction and β APP processing

Gang Yu^{*†}, Masaki Nishimura^{*†}, Shigeki Arawaka^{*}, Diane Levitan[‡], Lili Zhang[‡], Anurag Tandon^{*}, You-Qiang Song^{*}, Ekaterina Rogaeva^{*}, Fusheng Chen^{*}, Toshitaka Kawarai^{*}, Agnes Supala^{*}, Lyne Levesque^{*}, Haung Yu^{*}, Dun-Sheng Yang^{*}, Erin Holmes^{*}, Paul Milman^{*}, Yan Liang^{*}, Dong Mei Zhang^{*}, Dong Hong Xu^{*}, Christine Sato^{*}, Evgeny Rogaev^{*}, Marsha Smith[‡], Christopher Janus^{*}, Yanni Zhang[§], Ruedi Aebersold[§], Lindsay Farrer^{||}, Sandro Sorbi[¶], Amalia Bruni[#], Paul Fraser^{*} & Peter St George-Hyslop^{*}

^{*} Centre for Research in Neurodegenerative Diseases; The University Health Network (Toronto Western Hospital); and Departments of Medicine (Neurology) and Medical Biophysics, University of Toronto, Tanz Neuroscience Building, 6 Queen's Park Crescent West, Toronto, Ontario, M5S 3H2, Canada

[‡] Department of CNS and Cardiovascular Research, Schering Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, New Jersey 07033-0539, USA

[§] Molecular Biotechnology, University of Washington, Seattle, Washington 98195-7730, USA

^{||} Genetics Program, Boston University School of Medicine, 80 East Concord Street, Boston, Massachusetts 02118-2394, USA

[¶] Department of Neurology, University of Florence, viale Morgagni 65, Florence 50134, Italy

[#] Centro Regionale di Neurogenetica, USL 6, Lamezia Terme 88046, Italy

[†] These authors contributed equally to this work

Nicastrin, a transmembrane glycoprotein, forms high molecular weight complexes with presenilin 1 and presenilin 2. Suppression of nicastrin expression in *Caenorhabditis elegans* embryos induces a subset of *notch*/*glp-1* phenotypes similar to those induced by simultaneous null mutations in both presenilin homologues of *C. elegans* (*sel-12* and *hop-1*). Nicastrin also binds carboxy-terminal derivatives of β -amyloid precursor protein (β APP), and modulates the production of the amyloid β -peptide (A β) from these derivatives. Missense mutations in a conserved hydrophilic domain of nicastrin increase A β ₄₂ and A β ₄₀ peptide secretion. Deletions in this domain inhibit A β production. Nicastrin and presenilins are therefore likely to be functional components of a multimeric complex necessary for the intramembranous proteolysis of proteins such as Notch/GLP-1 and β APP.

Presenilin 1 (PS1)¹ and presenilin 2 (PS2)² are conserved, polytopic transmembrane proteins that are functional components of separate high molecular weight complexes^{3,4} resident in the endoplasmic reticulum (ER) and Golgi apparatus⁵. PS1 and PS2 are essential for an unusual, but poorly understood form of proteolytic cleavage that occurs within the transmembrane domains of several proteins, including β APP⁶, Notch^{7–10} and Irfp^{11,12}. In addition, the intramembranous proteolysis of β APP (termed γ -secretase cleavage) is increased by missense mutations in the presenilins associated with familial Alzheimer's disease (FAD)^{1,2}, resulting in overproduction of the neurotoxic A β derivative^{13,14}. Evidence suggests that the presenilins may have direct catalytic activity^{15,16}; however, three observations indicate that this activity requires at least interactions between the presenilins and other proteins. First, the abundance of PS1 and PS2 is tightly regulated by the limited abundance of another unknown component of the presenilin complexes¹⁷. Second, PS1 co-fractionates with γ -secretase enzymatic activity in a very high molecular weight complex¹⁸. Third, loss-of-function mutations in two intramembranous aspartate residues¹⁵ alter the structure and size of the presenilin complexes¹⁹, suggesting that these aspartate residues also affect critical interactions between the presenilins and other components of these complexes.

Here we report the identification of a new component of PS1 and PS2 complexes. This component, nicastrin, is a Type I transmembrane glycoprotein which interacts with both PS1 and PS2, and which has a central role in presenilin-mediated processing of β APP and some aspects of *notch*/*glp-1* signalling *in vivo*. The name nicastrin reflects the fact that the quest for the molecular machinery causing the presenilin-associated forms of Alzheimer's disease began nearly 40 years ago with the description of Alzheimer's disease in descendants of an extended family originating in the Italian village of Nicastro^{20,21}.

Isolation of nicastrin

HEK293 cells physiologically process β APP to A β peptide, contain β -secretase and display the same biochemical effects of mutations in Alzheimer's disease genes (β APP, PS1 and PS2) as observed *in vivo* in human and murine brain. We therefore used an anti-PS1 antibody (Ab1142 to residues 30–45) to immuno-extract PS1 and tightly associated proteins from intracellular membrane fractions of HEK293 cells expressing moderate levels of PS1. Coomassie-blue-stained SDS–polyacrylamide gel electrophoresis (PAGE) gels of the immuno-purified proteins contained two intense bands in addition to those of the PS1 holoprotein and its fragments (see Supplementary Information). Solid phase microextraction-capillary zone electrophoresis-microelectrospray (SPE-CE) tandem mass spectroscopy²² identified the proteins in these bands as a new protein (nicastrin), and α - and β -catenin. Catenins have previously been shown to interact with presenilins^{3,4,23,24}.

The deduced amino-acid sequence of the nicastrin peptide was identical to that predicted for several anonymous partial complementary DNAs in public databases. The partial cDNAs were used to derive a full-length nicastrin cDNA (GenBank Accession number AF240468). The human *nicastrin* gene maps to a region of chromosome 1 (near *D1S1595* and *D1S2844*) that, in two independent genome-wide surveys^{25,26}, has generated evidence for genetic linkage and/or allelic association with an Alzheimer's disease susceptibility locus. The *nicastrin* gene encodes an open reading frame of 709 amino acids containing a putative amino-terminal signal peptide; a long N-terminal hydrophilic domain containing glycosylation, N-myristoylation and phosphorylation motifs; a ~20-residue hydrophobic putative transmembrane domain; and a short hydrophilic carboxy terminus of 20 residues (Fig. 1). However, we could not find any significant amino-acid sequence homology or strong motif similarity to other functionally characterized proteins.

In the absence of homology to other proteins, we screened sequence databases for orthologous genes in other species. We found a full-length *C. elegans* nicastrin orthologue (ZC434.6) in public databases (accession no. Q23316; $P = 1e^{-37}$; identity = 22%; similarity = 41%). We cloned and sequenced full-length murine and *Drosophila* nicastrin orthologues from appropriate cDNA libraries

using partial cDNA sequences from these databases as start points (mouse nicastrin accession no. AF24069, $P = 0.00$, identity = 89%, similarity = 93%; *D. melanogaster* nicastrin accession no. AF240470, $P = 4e^{-77}$, identity = 30%, similarity = 48%). A weakly similar protein is also predicted to exist in the plant *Arabidopsis* (accession no. CAB89225.1, $P = 8e^{-22}$, identity = 26%, similarity = 40%). The four animal nicastrins have similar predicted topologies and have three domains with significant sequence conservation near residues 306–360, 419–458, and 625–662 of human nicastrin ($P < 1e^{-6}$; identities = 50–89%, similarities = 60–90% for each segment) (Fig. 1). Within the first conserved domain, all four proteins contained the motif DYIGS (residues 336–340), which is also partially conserved in the *Arabidopsis* protein. All four animal nicastrins also contain four cysteines spaced at 16–17-residue intervals in the N terminus (Cys 195, Cys 213, Cys 230 and Cys 248).

Nicastrin interacts with presenilins

Western blots of lysates from HEK293 cells transfected with a V5-epitope-tagged nicastrin cDNA revealed a V5-immunoreactive band with a relative molecular mass of about 110,000 ($M_r \approx 110K$). Following digestion by Endo H, this band was reduced to ~80K (the predicted size of the nicastrin amino-acid sequence), confirming that nicastrin is glycosylated (Fig. 2, lanes 1–4). The V5-immunoreactive band was also detected by the anti-nicastrin antibodies anti-N-NCT (against human nicastrin residues 290–406) and anti-C-NCT (against residues 691–709) (Fig. 2, lane 5). In cells transiently transfected with V5-nicastrin, a series of $M_r \sim 7K$ –10K fragments were observed which we predict contain the transmembrane domain and short C terminus of nicastrin plus the V5 epitope of M_r 3K (Fig. 2, lane 1). However, because they are not detectable in stably transfected cells (Fig. 2, lane 3) and no secreted N-terminal fragments were detected in conditioned media, it is unclear whether or not these C-terminal derivatives are artefacts of over-expression.

Immunoprecipitation studies in human brain (Fig. 3a) and in transfected HEK293 cells (see Supplementary Information) confirm

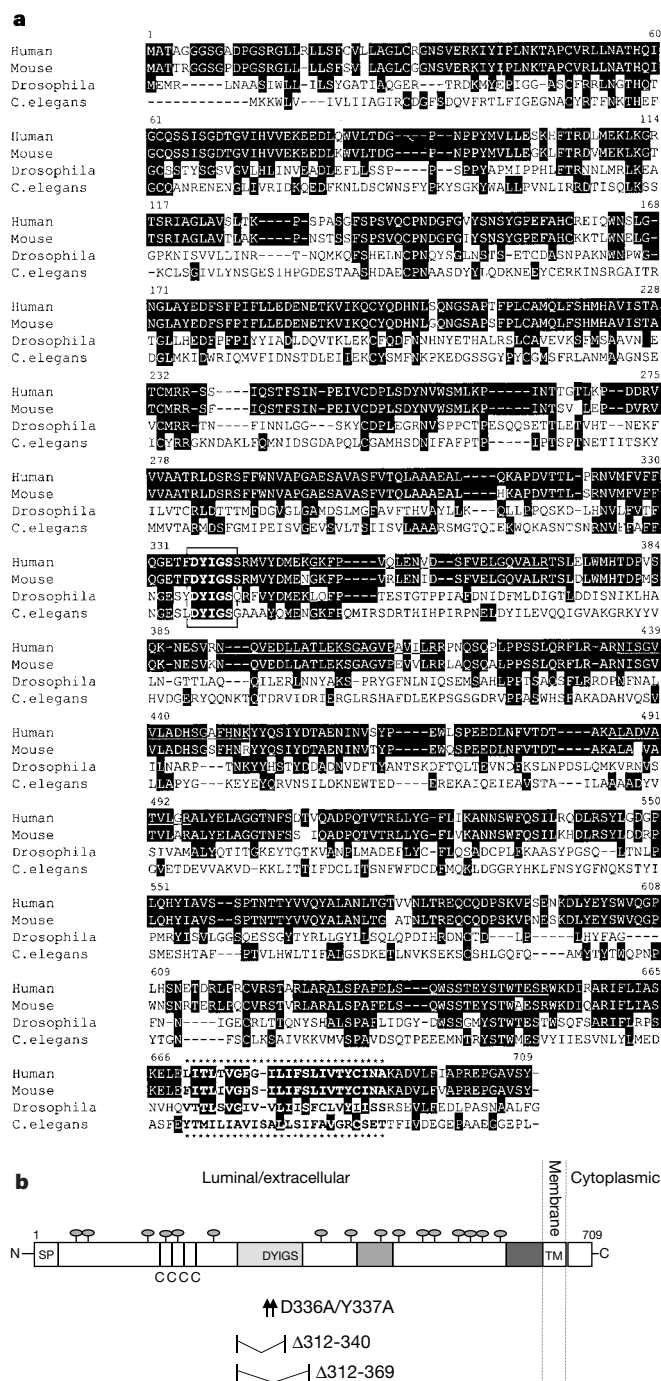


Figure 1 Predicted amino-acid sequence and topology of nicastrin. **a**, Predicted sequence of human nicastrin and murine, *D. melanogaster* and *C. elegans* nicastrin orthologues. Asterisks denote a putative transmembrane domain. Peptide fragments identified by mass spectrometry are underlined. Conserved sequences are highlighted. The DYIGS motif is boxed. **b**, Predicted topology of human, murine, *D. melanogaster* and *C. elegans* nicastrin orthologues. SP, signal peptide; C, four conserved cysteines; ovals, potential glycosylation sites; TM, transmembrane domain. Shading indicates highly conserved domains. The location of the DYIGS motif and functionally active missense (arrows) and deletion mutations (between vertical lines) are shown.

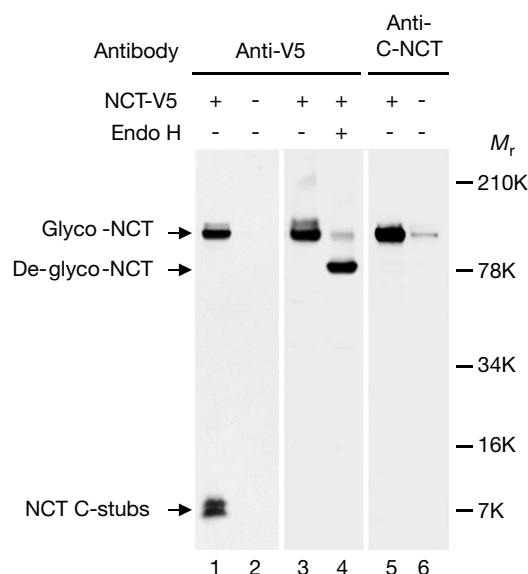


Figure 2 Characterization of nicastrin (NCT) using V5-tagged nicastrin. Immunoblots of lysates from HEK293 cells either transiently transfected (lane 1) or stably transfected with NCT-V5 (lanes 3–5) probed with anti-V5 antibody (lanes 1–4) or with a C-terminal anti-NCT antibody (anti-C-NCT, lanes 5–6). In transiently transfected cells, C-terminal stubs of nicastrin can be detected, which are not present in stable cell lines. Endo H treatment shows nicastrin is glycosylated (lane 4). The anti-C-NCT antibody (lane 5) detects the same band as the anti-V5 antibody in NCT-V5 stably transfected cells, and also detects a small amount of endogenous nicastrin in untransfected cells (lane 6).

that the nicastrin–PS1 interaction is authentic and specific ($n > 4$ replications). Thus, pre-immune serum, antibodies to other proteins resident in the endoplasmic reticulum (for example, calnexin), and other V5-tagged proteins (in studies using HEK293 cells) do not co-precipitate either nicastrin or PS1. Glycerol velocity centrifugation gradients showed that nicastrin co-fractionates with the high molecular weight PS1 and PS2 complexes. Nicastrin and PS1 strongly co-localize with markers of the endoplasmic reticulum and Golgi in MDCK cells, and their patterns of transcription overlap on northern blots from human tissues (see Supplementary Information). Parallel studies confirmed that nicastrin also interacts with PS2 in human brain (Fig. 3a) and in HEK293 cells (see

Supplementary Information). Finally, co-immunoprecipitation experiments revealed that nicastrin interacts equivalently with wild-type PS1, with PS1-L392V (an FAD-related gain-of-function mutant that increases A β secretion^{1,14}) and with PS1-D385A (a loss-of-function mutant that inhibits γ -secretase and A β secretion¹⁵) ($n = 3$ replications, Fig. 3b).

Nicastrin and *notch* signalling

To explore whether nicastrin, like the presenilins, might have a role in *notch* signalling *in vivo*, we used RNA interference (RNAi) in *C. elegans*. Wild-type worms injected with *C. elegans* nicastrin double-stranded (ds) RNA produced dead embryos, many of which lacked an anterior pharynx (Fig. 4, bottom panel). This phenotype was highly reproducible ($n > 35$ animals injected in 5 replications) and specific (the phenotype was not observed with injection buffer or with ~ 50 unrelated double stranded RNAs). Except for embryonic lethality, none of the other phenotypes associated with a lack of *C. elegans* presenilin (*sel-12*) activity were observed. However, this phenotype is identical to that induced when the activity of genes in the *notch/glp-1* pathway (*glp-1*, *aph-1* or *aph-2*) are reduced, or when the activities of both *C. elegans* presenilin homologues (*sel-12* and *hop-1*) are reduced simultaneously^{27–29}. Subsequently, we learned that *aph-2* has been mapped to a spontaneous transposon insertion mutant in the ZC434.6 gene³⁰, confirming that *aph-2* and nicastrin are identical and that nicastrin contributes to some aspects of *notch/glp-1* signalling in *C. elegans* embryos.

Nicastrin modulates γ -secretase cleavage of β APP

The β APP holoprotein (FL- β APP) is physiologically processed by a two-step proteolytic pathway. Initially, FL- β APP is cleaved near the cell surface in its extracellular domain either by β -secretase (to generate a membrane-bound 99-residue C-terminal stub, C99- β APP)^{31–34}, or by a putative α -secretase (to generate an 83-residue membrane-bound C-terminal stub, C83- β APP). These stubs are then further cleaved within their transmembrane domains by the presenilin-linked γ -secretase to generate A β and p3 respectively. The γ -site cleavage can occur at either of two positions to form

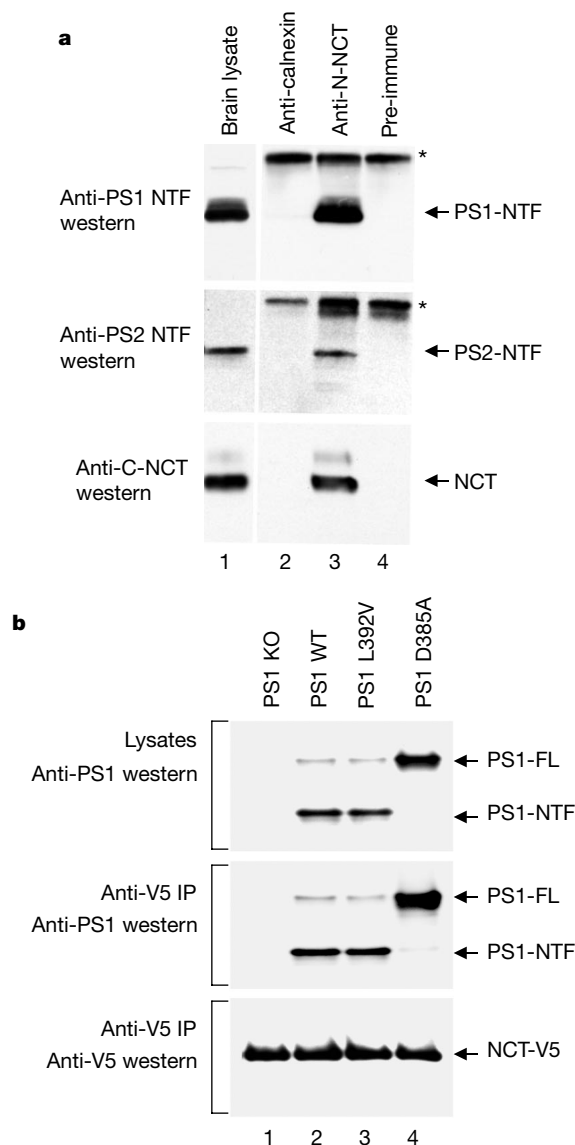


Figure 3 PS1 and PS2 form complexes with nicastrin. **a**, In digitonin-solubilized human brain lysates, antibody to the N terminus of nicastrin (anti-NCT; lane 3) co-precipitates endogenous PS1 (top panel), PS2 (middle panel) and nicastrin itself (bottom panel). Antibodies to calnexin (lane 2) and pre-immune serum (lane 4) do not co-precipitate PS1, PS2 or nicastrin. Asterisk corresponds to immunoglobulin. The column labelled 'brain lysate' (lane 1) contains about 25% of the starting lysate used for immunoprecipitation. **b**, Nicastrin interacts equivalently with mutant and wild-type PS1. Anti-V5 immunoprecipitation products from NCT–V5 transiently transfected embryonic fibroblasts from PS1-knockout mice or from wild-type mice stably overexpressing human wild-type PS1 (PS1-WT), PS1-L392V or PS1-D385A were investigated for PS1 with anti-PS1 antibody 14.3 (middle panel), or for NCT–V5 (bottom panel). Western blot of the starting lysates (top panel) shows the initial amounts of PS1.

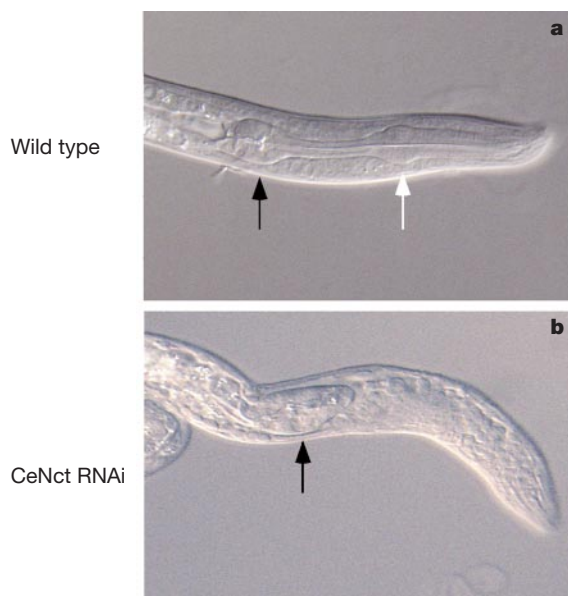


Figure 4 Effect of nicastrin on Notch signalling. Black arrow points to the posterior bulb of the pharynx; white arrow points to anterior bulb of the pharynx. **a**, Newly hatched non-injected L1 animal. **b**, Newly hatched L1 animal after dsRNA injection into a wild-type parent. Note the complete absence of the anterior bulb of the pharynx in nicastrin RNAi-treated *C. elegans* embryos.

A β_{40} (predominant isoform) or A β_{42} (fibrillogenic, neurotoxic isoform)³⁵.

To assess the role of nicastrin–presenilin complexes in β APP processing, we investigated anti-nicastrin co-immunoprecipitation products to determine whether nicastrin might directly interact with β APP and/or its C-terminal derivatives. An antibody to the C terminus of β APP (Ab369) detected both FL- β APP and C99- β APP/C83- β APP in the anti-nicastrin co-immunoprecipitation products from digitonin lysates of HEK293 cell lines stably expressing wild-type β APP ($n = 4$ replications, Fig. 5a). This was confirmed in HEK293 cells over-expressing either β APP_{Swedish} or the SpC99- β APP cDNA (encoding the signal peptide and C-terminal 99 residues of β APP) ($n > 4$ replications, Fig. 5b). In Fig. 5b, nicastrin appears to interact better with C99- β APP than with C83- β APP; however, C83- β APP is less abundant in these cells (Fig. 5b, lanes 1–4). In fact, nicastrin interacts with both C99- β APP and C83- β APP stubs (Fig. 5c, lane 9). Of greater interest, the genotype of the co-expressed PS1/PS2 molecule dynamically influenced the interaction of nicastrin with C99- β APP/C83- β APP. After transient transfection of nicastrin, more C99- β APP/C83- β APP stubs

co-immunoprecipitated with nicastrin in cells expressing the FAD-associated PS1-L392V mutation than in cells expressing wild-type PS1 (and equivalent quantities of nicastrin and C99- β APP) ($n = 4$ replications, Fig. 5b, lanes 2 and 4). Conversely, much less C99- β APP/C83- β APP stubs co-immunoprecipitated with nicastrin in cell lines expressing the loss-of-function PS1-D385A mutation (despite the presence of very large amounts of C-terminal β APP derivatives in these cells) ($n = 4$ replications, Fig. 5b, lanes 5 and 6).

Similar, but less robust effects were also observed in cells over-expressing full-length β APP_{Swedish} plus PS2 (wild-type, FAD PS2-N141I or loss-of-function PS2-D366A mutants) (Fig. 5b, lanes 9–11). More notably, the PS1-sequence-related differences in the nicastrin–C99- β APP/C83- β APP interaction were most evident within 24 h of transfection with nicastrin (Fig. 5b, c). By 72 h, the amount of C99- β APP/C83- β APP that immunoprecipitated with nicastrin reached a new equilibrium, and was roughly equivalent in cells expressing different PS1-sequences ($n = 2$ replications, Fig. 5c). This dynamic change was not due to changes in the levels of PS1, C-terminal β APP derivatives (data not shown) or nicastrin (Fig. 5c). One interpretation is that the presenilins may be

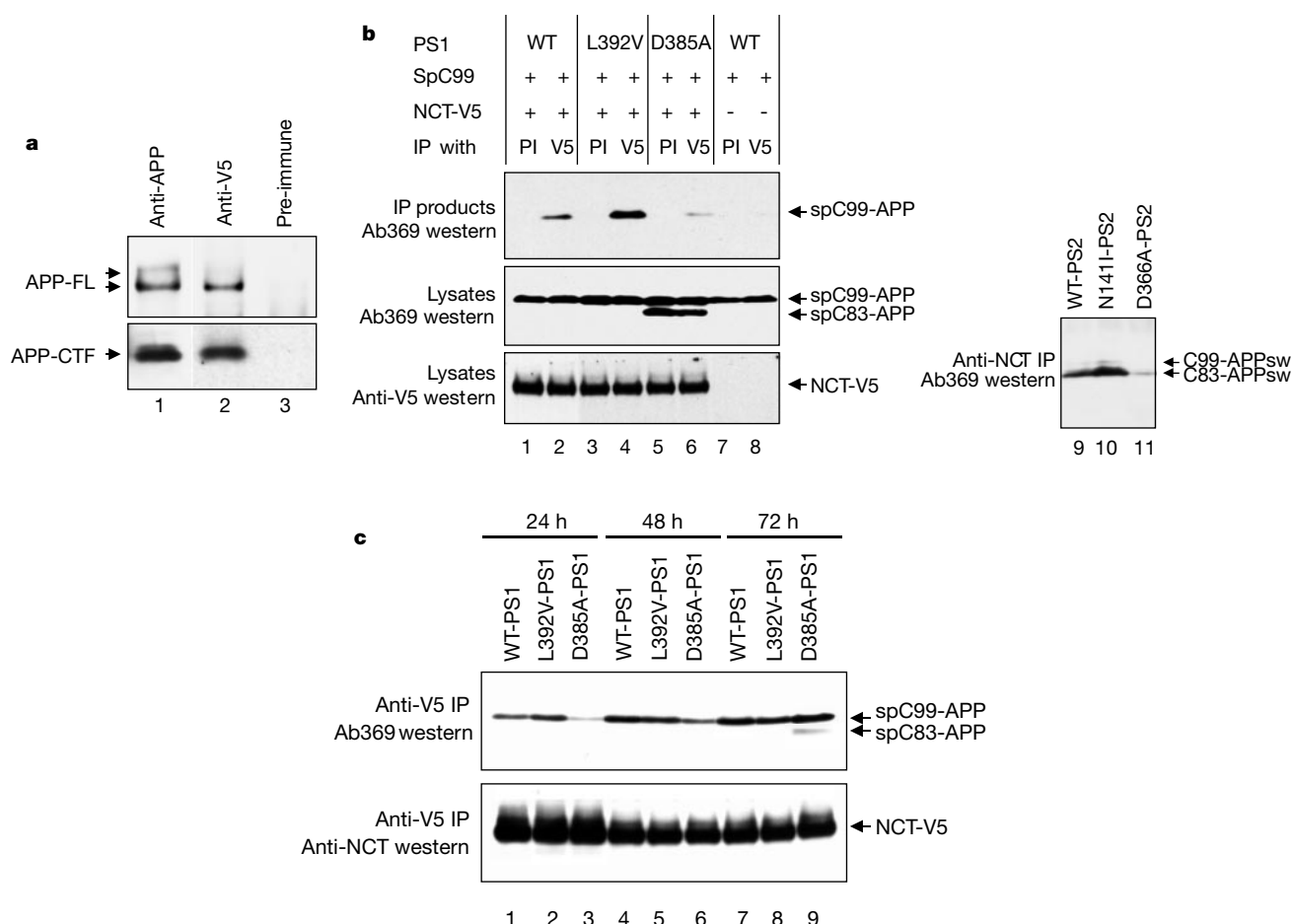


Figure 5 α -secretase and β -secretase cleavage fragments of β APP (C83- β APP and C99- β APP respectively) co-precipitate with nicastrin. **a**, Western blot of anti-NCT-V5 immunoprecipitates from HEK293 cells expressing nicastrin and wild-type β APP reveals FL- β APP and C99- β APP/C83- β APP (detected with antibody 369; lane 2). FL- β APP and C99- β APP/C83- β APP do not co-precipitate with pre-immune serum (lane 3). Note, for anti-APP (lane 1), about 10% of the starting lysate was used for immunoprecipitation. **b**, Top left, HEK293 cells stably expressing the SpC99- β APP construct together with either wild-type PS1 (WT-PS1), L392V-PS1 or D385A-PS1 were transiently transfected with NCT-V5 for 24 h. The NCT-V5 immunoprecipitation products ('V5' lanes) but not the pre-immune immunoprecipitation products (PI) contain C99- β APP fragments (detected with antibody 369 to C terminus of β APP). Middle left, anti-C-terminus of β APP

immunoblot of the same lysates (antibody 369) showing starting levels of C99- β APP/C83- β APP (increased amounts of C83- β APP in lanes 5 and 6 reflect inhibition of γ -secretase cleavage of C83- β APP by the D385A-PS1 mutant). Bottom left, anti-V5 immunoblot of the same lysates showing similar starting levels of NCT-V5. Bottom right, identical results were obtained with cells stably expressing PS2 (wild type, N141I, D366A) and full-length APP (APP_{sw}), and transiently transfected with untagged nicastrin for 24 h. **c**, In the same HEK293 lines, the genotype at PS1 dynamically influences the amount of β APP stubs co-immunoprecipitating with nicastrin. The differences are most marked 24 h after transient transfection of nicastrin, and reach a new higher, steady state at 72 h. Note that the D385A-PS1 cells have higher basal levels of C99- β APP and C83- β APP fragments compared with wild-type cells.

dynamically involved in regulating or loading nicastrin with the substrates of γ -secretase.

To explore the role of nicastrin in $A\beta$ production, we created a series of HEK293 cell lines stably overexpressing $\beta APP_{\text{Swedish}}$ plus either wild-type or mutant nicastrin. Relative to mock-transfected or LacZ-transfected cells, overexpression of wild-type and most mutant nicastrins had no significant effect on $A\beta$ secretion. However, missense mutation of the conserved DYIGS motif to AAIGS (residues 336–340) caused a significant increase in $A\beta$ secretion, and especially in $A\beta_{42}$ secretion ($P < 0.001$, Fig. 6; 5 independent clonal cell lines, 29 total observations). Conversely, deletion of the DYIGS domain in two independent

constructs (NCT Δ 312–369 and NCT Δ 312–340) caused a significant reduction in both $A\beta_{42}$ and $A\beta_{40}$ secretion. This was more profound in NCT Δ 312–369 cells than in NCT Δ 312–340 cells (Fig. 6, five independent clones in total). These effects are similar to the effects of gain-of-function and loss-of-function mutations in the presenilins, respectively. However, in contrast to PS1^{−/−} and PS1-D385A mutants, the reduction in $A\beta$ secretion induced by the NCT Δ 312–369 and NCT Δ 312–340 mutants was not accompanied by accumulation of C99- βAPP and C83- βAPP stubs. This suggests that C99- βAPP and C83- βAPP stubs that do not enter the nicastrin–presenilin complex for γ -secretase cleavage are degraded by other pathways.

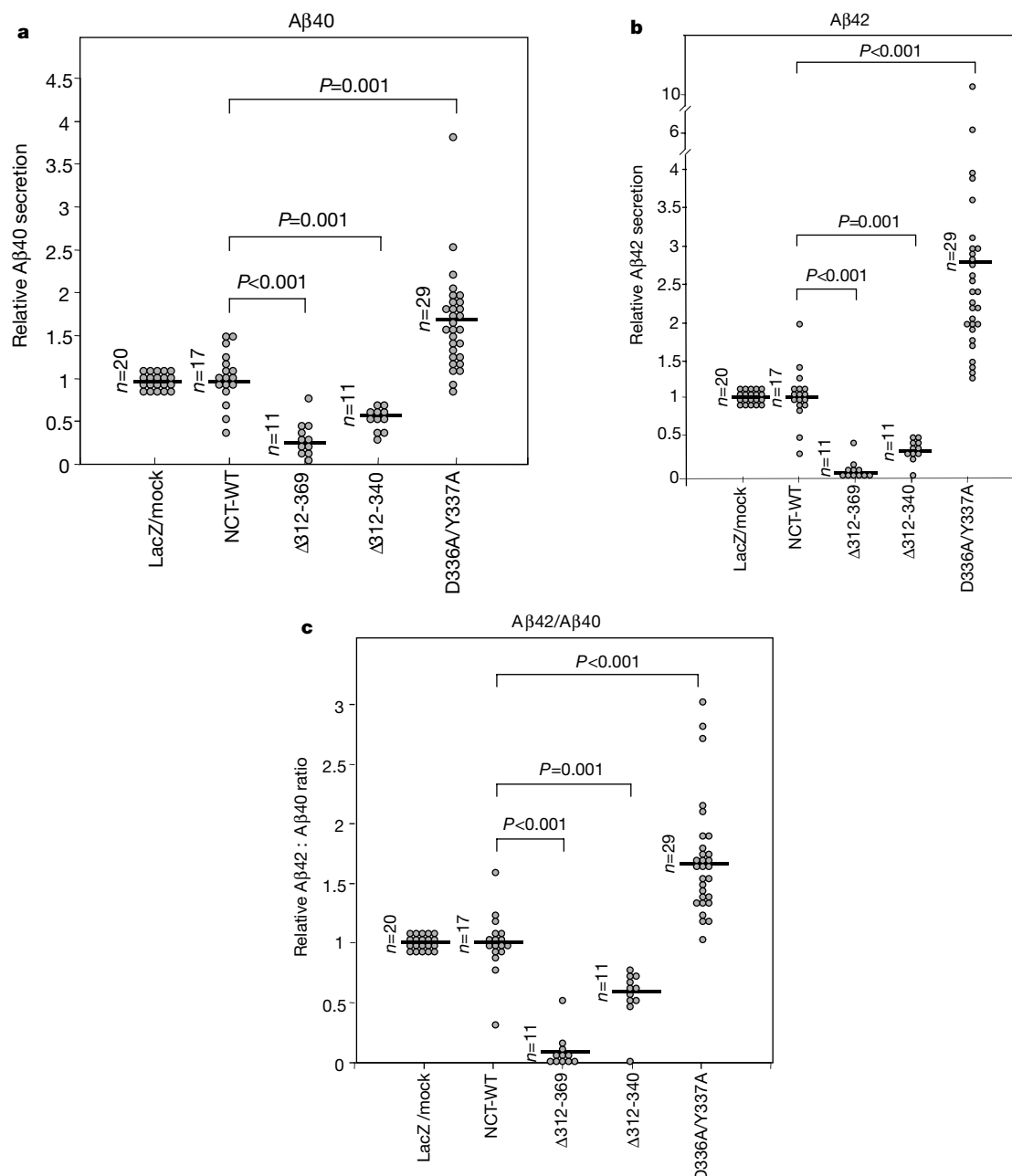


Figure 6 Scatter plots of secreted $A\beta_{40}$, $A\beta_{42}$ and $A\beta_{42}/A\beta_{40}$ ratios showing increased $A\beta$ secretion from HEK293 cells expressing DYIGS→AAIGS mutant nicastrin, and decreased $A\beta$ secretion from cells expressing Δ 312–369 or Δ 312–340 mutants. Each dot represents a single experiment performed on LacZ/mock transfected controls expressing APP_{sw} (lane 1, 3 independent cell lines); wild-type nicastrin plus APP_{sw} (lane 2, 5 cell lines); Δ 312–369 NCT plus APP_{sw} (lane 3, 2 cell lines); Δ 312–340 NCT plus APP_{sw}

(lane 4, 3 cell lines); DYIGS→AAIGS NCT plus APP_{sw} (lane 5, 5 cell lines). The $A\beta$ level in mock-transfected or LacZ-transfected control samples was normalized to 1.0. $A\beta$ levels in test samples were expressed relative to this normalized value. Mean normalized values are depicted by horizontal bar (see Supplementary Information). P values represent comparisons of normalized $A\beta$ levels from media of cells expressing mutant nicastrin relative to levels from cells expressing wild-type nicastrin.

The effects of nicastrin mutations on A β secretion are not due to trivial explanations such as differences in the levels of nicastrin, β APP holoprotein or PS1/PS2, nor are they due to effects on the activity of α - or β -secretase (see Supplementary Information). None of the mutations affect the interaction of nicastrin with C99-/C83- β APP (Fig. 7). However, although the DYIGS $\rightarrow$ AAIGS mutant of nicastrin can efficiently co-immunoprecipitate PS1, both of the deletion mutants significantly reduced the nicastrin-PS1 interaction (Fig. 7, two independent cell lines each, in triplicate). Furthermore, the magnitude of this effect was proportional to the effect of each of the deletion mutants on A β secretion (Fig. 7). Residues 312–369 of nicastrin contain no obvious functional motifs or sequence homology to other known proteins. Consequently, mutations in this domain presumably either affect a presenilin-binding domain, or they affect a regulatory domain that modulates both the nicastrin-PS1 interaction and the subsequent γ -secretase-mediated cleavage of nicastrin-bound C99- β APP and C83- β APP stubs.

Conclusions

Our data indicate that nicastrin is an authentic, functional component of PS1 and PS2 complexes. Nicastrin is unlikely to be simply another Type I membrane protein whose processing involves the presenilins (for example, like β APP). First, nicastrin is a principal stoichiometric component of the presenilin complexes. In contrast, substrates like β APP and Notch are not principal constituents of the high molecular weight presenilin complexes, and can only be inconsistently co-precipitated with PS1/PS2 (refs 3, 36–39). Second, the role of nicastrin in the presenilin-mediated processing of both β APP and Notch would not be expected if nicastrin were simply another substrate.

These results suggest that nicastrin is part of a new functional complex (putatively a 'secretosome') involved in the unusual intramembranous proteolytic processing of transmembrane proteins like β APP and Notch. The exact role of nicastrin is currently unclear. Although the primary amino-acid sequence of nicastrin does not resemble known proteases, we cannot entirely exclude a

proteolytic function. However, the current data are compatible with two other models. One model for nicastrin activity is a role in binding the substrates of the presenilin- γ -secretase complexes (and/or other γ -secretase-like enzymes). Our data suggest that this binding is dynamically modulated by presenilin mutations in a direction commensurate with the effect of each particular presenilin mutation on γ -secretase activity. Thus, presenilin FAD-mutations increase nicastrin binding to C99- β APP/C83- β APP and increase A β secretion whereas presenilin aspartate mutants have the opposite effect.

The alternative model for nicastrin function is that it might regulate the activity of γ -secretase. This second model is supported by the observation that mutation of conserved residues in the N-terminal domain of nicastrin can selectively increase or decrease A β production. This domain is likely to be located in the lumen of intracellular membrane-bound organelles (our own unpublished data). Binding of putative ligands to this domain, as part of a sensor mechanism, might allow nicastrin to regulate γ -secretase activity perhaps in response to conditions of β APP processing in the secretory pathway. In this second model, missense mutation of the DYIGS domain would cause a constitutive gain-of-function effect in which nicastrin overactivates γ -secretase cleavage. Conversely, the deletion mutants still bind C99- β APP/C83- β APP, but might prevent these deletion-mutant nicastrins from activating γ -secretase cleavage because of a reduced interaction with the presenilins. Experiments to mix gain-of-function PS1/PS2 mutants with loss-of-function nicastrin mutants (and vice versa) will be needed to dissect the functional order of these molecules. However, β APP is upstream of both PS1/PS2 and nicastrin because loss-of-function and/or null mutations in either nicastrin or the presenilins block the effects of the β APP_{Swedish} mutation.

Whether or not genetic variants in nicastrin are associated with inherited susceptibility to Alzheimer's disease remains to be seen. In preliminary studies, we have found no mutations or polymorphisms in the open reading frame of nicastrin in affected members of 19 late onset FAD pedigrees in which no obligate recombinants were detected between Alzheimer's disease and the *DIS1595*–*14 cM-DIS2844* genetic interval containing nicastrin. However, these results do not preclude the existence of risk alleles in other data sets and/or in non-coding regulatory sequences of nicastrin in families within our data set. Furthermore, these data do not abrogate a role for wild-type nicastrin in the pathogenesis of abnormal β APP processing in Alzheimer's disease due to environmental and other genetic causes. Indeed, because manipulation of an exposed domain of nicastrin has significant effects on β APP processing, nicastrin is potentially a tractable target for therapeutic modulation of A β production in patients with Alzheimer's disease and related disorders. □

Methods

Purification of the nicastrin protein

Membrane proteins were purified from HEK293 cells stably overexpressing moderate levels of wild-type human PS1, extracted with digitonin and fractionated on a 10–40% glycerol gradient containing 0.5% digitonin as described^{3,40}. The peak PS1-containing fractions (which contain ~0.5 g of membrane proteins) were pooled and incubated overnight with Protein A/G agarose coupled with either an affinity-purified PS1 NTF antibody (antibody 1142; P. E. Fraser, unpublished data) or a control IgG purified from pre-immune rabbit serum. After washing six times with buffer (1% digitonin, 0.5% CHAPS, 20 mM HEPES pH 7.2, 100 mM KCl, 10 mM CaCl₂, 5 mM MgCl₂ and a protease inhibitor cocktail), isolated protein complexes were eluted with 0.1 M glycine-HCl pH 3.0, separated by Tris-glycine SDS-PAGE gels, and stained with silver stain or Coomassie blue stain. Individual protein bands were cut out and analysed with SPE-CE-tandem mass spectroscopy²². Protein bands were first digested in-gel with trypsin; the digested peptides were extracted and concentrated in a speed vacuum; the peptides were separated by solid-phase extraction capillary electrophoresis and analysed by on-line tandem mass spectrometry. High quality MS/MS spectra were selected and used for the protein sequence and translated nucleotide sequence database searches.

Nicastrin cDNAs

Full-length human, murine and *Drosophila melanogaster* nicastrin cDNAs were obtained

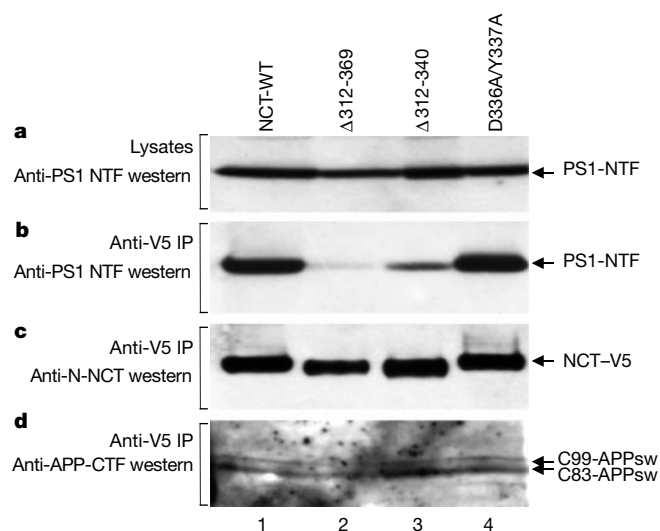


Figure 7 Functional nicastrin mutants do not significantly impair the nicastrin interaction with C99- β APP/C83- β APP in HEK293 cells expressing β APP_{Swedish}. The DYIGS $\rightarrow$ AAIGS mutant of NCT has no discernible effect on the NCT-PS1 interaction, but the NCT Δ 312–369 or NCT Δ 312–340 mutations do impair the NCT-PS1 interaction proportional to their effect on A β secretion. **a**, Lysates probed with anti-PS1-NTF antibody (Ab14). **b**, Anti-V5 NCT immunoprecipitation products probed with anti-PS1-NTF antibody (Ab14). **c**, Anti-V5 NCT immunoprecipitation products probed with anti-NCT antibody (anti-N-NCT). **d**, Anti-V5 NCT immunoprecipitation products probed with anti-C-terminal β APP antibody 369. Note, a longer exposure was used for this immunoblot.

using oligonucleotides designed from partial cDNA/expressed sequence tag sequences in public databases (<http://www.ncbi.nlm.nih.gov>) to screen appropriate cDNA libraries, for 5' RACE and/or for polymerase chain reaction with reverse transcription (RT-PCR) experiments. Chromosomal locations and genetic map positions of the murine and human nicastrins were obtained from public genetic and transcriptional maps (<http://www.ncbi.nlm.nih.gov>). A nicastrin expression construct was generated by inserting human nicastrin cDNA in-frame with the V5 epitope of pcDNA6 at the C terminus of nicastrin.

Cell lines

HEK293 cells were transiently or stably transfected with either V5-tagged nicastrin or V5-LacZ or empty plasmid (controls). We have noticed that the level of nicastrin in cells stably expressing nicastrin constructs tends to fall over time, suggesting that overexpression of nicastrin may be slightly toxic. We also used previously characterized HEK293 cell lines expressing PS1/PS2 and/or β APP or SP-C99- β APP to evaluate the interaction of nicastrin with PS1, PS2 and β APP^{34,41,42}.

Biochemical methods

Brain tissue, cultured cells and partially purified cell membranes were lysed with 1% digitonin lysis buffer or with 1% NP40, and the protein extracts were subjected to gradient fraction, immunoprecipitation or direct western blotting as described³.

Antibodies

V5-epitope-tagged nicastrin was detected using anti-V5 (Invitrogen, CA). Polyclonal rabbit anti-nicastrin anti-sera were raised against residues 290–406 of human nicastrin fused to glutathione S-transferase (anti-N-NCT), or against a peptide containing residues 691–709 (anti-C-NCT). Antibodies included anti-PS1-NCT (Ab14, S. Gandy; 1142); or anti-PS1-CTF (N. Brindle); anti-PS2 (DT2, P. Davies); anti-FL- β APP and C-terminal α - and β -secretase derivatives (369, S. Gandy).

A β assays

A β ₄₀ and A β ₄₂ levels were measured as described⁴³ by ELISA using 6–20-h conditioned media collected from HEK293 cells stably expressing both β APP (β APP^{Swedish} or β APP^{wild-type}) and either nicastrin (wild type or one of its mutants) or a control sequence (LacZ or empty vector). The A β levels in cells transfected with a control sequence (LacZ or empty vector) were set at 1.0 and the A β levels in test samples were then normalized to this value. Transformed (square root transformation) data were compared by a priori planned pairwise Student's *t*-test⁴⁴.

RNA interference

Sense and antisense RNA were transcribed *in vitro* from PCR products amplified from the *C. elegans* nicastrin cDNA phage yk477b8 (Y. Kohara). After annealing equal quantities of sense and antisense products, we injected the dsRNA into L4 stage animals. Injected animals were transferred to fresh plates daily and the progeny scored at least 36 h after injection for the embryonic lethal phenotype and for 4–5 days after injection for other phenotypes. The penetrance of the embryonic lethal phenotype ranged from 75 to 100% of the total brood.

Received 24 March; accepted 21 July 2000.

- Sherrington, R. *et al.* Cloning of a gene bearing missense mutations in early onset familial Alzheimer's disease. *Nature* **375**, 754–760 (1995).
- Rogaev, E. I. *et al.* Familial Alzheimer's disease in kindreds with missense mutations in a novel gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* **376**, 775–778 (1995).
- Yu, G. *et al.* The presenilin 1 protein is a component of a high molecular weight intracellular complex that contains β -catenin. *J. Biol. Chem.* **273**, 16470–16475 (1998).
- Nishimura, M. *et al.* Presenilin mutations dominantly modulate β -catenin trafficking. *Nature Med.* **5**, 164–169 (1999).
- De Strooper, B. *et al.* Post-translational modification, subcellular localization and membrane orientation of the Alzheimer's Disease associated Presenilins. *J. Biol. Chem.* **272**, 3590–3598 (1997).
- De Strooper, B. *et al.* Deficiency of presenilin 1 inhibits the normal cleavage of amyloid precursor protein. *Nature* **391**, 387–390 (1998).
- Ye, Y., Lukinova, N. & Fortini, M. E. Neurogenic phenotypes and altered Notch processing in *Drosophila* presenilin mutants. *Nature* **398**, 525–529 (1999).
- Struhl, G. & Greenwald, I. Presenilin is required for activity and nuclear access of Notch in *Drosophila*. *Nature* **398**, 522–525 (1999).
- De Strooper, B. *et al.* A presenilin dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–522 (1999).
- Song, W. *et al.* Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. *Proc. Natl Acad. Sci. USA* **96**, 6959–6963 (1999).
- Katayama, T. *et al.* Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nature Cell Biol.* **1**, 479–485 (1999).
- Niwa, M., Sidrauski, C., Kaufman, R. J. & Walter, P. A role for presenilin-1 in nuclear accumulation of Ire1 fragments and induction of the mammalian Unfolded Protein Response. *Cell* **99**, 691–702 (1999).
- Scheuner, D. *et al.* Secreted amyloid- β protein similar to that in the senile plaques of Alzheimer Disease is increased *in vivo* by presenilin 1 and 2 and APP mutations linked to FAD. *Nature Med.* **2**, 864–870 (1996).
- Citron, M. *et al.* Mutant presenilins of Alzheimer's Disease increase production of 42 residue amyloid β -protein in both transfected cells and transgenic mice. *Nature Med.* **3**, 67–72 (1997).

- Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin 1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* **398**, 513–517 (1999).
- Li, Y.-M. *et al.* Photoactivated gamma-secretase inhibitors directed to the active site covalently label presenilin 1. *Nature* **405**, 689–694 (2000).
- Thinakaran, G. *et al.* Evidence that levels of presenilins (PS1 and PS2) are coordinately regulated by competition for limiting cellular factors. *J. Biol. Chem.* **272**, 28415–28422 (1997).
- Li, Y.-M. *et al.* Presenilin 1 is linked with gamma-secretase activity in the detergent solubilized state. *Proc. Natl Acad. Sci. USA* **97**, 6138–6143 (2000).
- Yu, G. Mutation of conserved aspartates affect maturation of both aspartate-mutant and endogenous presenilin 1 and presenilin 2 Complexes. *J. Biol. Chem.* (in the press).
- Feldman, R. G., Chandler, K. A., Levy, L. L. & Glaser, G. H. Familial Alzheimer's Disease. *Neurology* **13**, 811–824 (1963).
- Foncin, J.-F. *et al.* Démence présénile d'Alzheimer transmise dans une famille étendue. *Rev. Neurol. (Paris)* **141**, 194–202 (1985).
- Figgeys, D., Ducret, A., Yates, J. R. & Aebersold, R. Protein identification by solid phase micro-extraction-capillary zone electrophoresis-microelectrospray-tandem mass spectroscopy. *Nature Biotech.* **14**, 1579–1583 (1996).
- Zhou, J. *et al.* Presenilin 1 interacts with a novel member of the armadillo family. *NeuroReport* **8**, 2085–2090 (1997).
- Levesque, G. *et al.* Presenilins interact with armadillo proteins including neural specific plakophilin related protein and beta-catenin. *J. Neurochem.* **72**, 999–1008 (1999).
- Kehoe, P. *et al.* A full genome scan for late onset Alzheimer's Disease. *Hum. Mol. Genet.* **8**, 237–245 (1999).
- Zubenko, G. S., Hughes, H. B., Stiffler, J. S., Hurtt, M. R. & Kaplan, B. B. A genome survey for novel Alzheimer disease risk loci: results at 10 cM resolution. *Genomics* **50**, 121–128 (1998).
- Preiss, J. R., Schnabel, H. & Schnabel, R. The glp-1 locus and cellular interactions in early *C. elegans* embryos. *Cell* **51**, 601–611 (1987).
- Li, X. & Greenwald, I. HOP-1, a *Caenorhabditis elegans* presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signalling. *Proc. Natl Acad. Sci. USA* **94**, 12204–12209 (1998).
- Westlund, B., Barry, D., Clover, R., Basson, M. & Johnson, C. D. Reverse genetic analysis of *C. elegans* presenilins reveals redundant but unequal roles for sel-12 and hop-1 in Notch-pathway signaling. *Proc. Natl Acad. Sci. USA* **96**, 2497–2502 (1999).
- Goutte, C., Hepler, W., Mickey, K. M. & Priess, J. R. aph-2 encodes a novel extracellular protein required for GLP-1-mediated signaling. *Development* **127**, 2481–2492 (2000).
- Hussain, I. *et al.* Identification of a novel aspartic protease (Asp 2) as beta-secretase. *Mol. Cell Neurosci.* **14**, 419–427 (1999).
- Sinha, S. *et al.* Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature* **402**, 537–540 (1999).
- Vassar, R. *et al.* Beta-secretase cleavage of Alzheimer's Amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741 (1999).
- Yan, R. *et al.* Membrane-anchored aspartyl protease with Alzheimer's Disease beta-secretase activity. *Nature* **402**, 533–537 (1999).
- Selkoe, D. J. Normal and abnormal biology of β -Amyloid Precursor Protein. *Ann. Rev. Neurosci.* **17**, 489–517 (1994).
- Xia, W., Zhang, J., Perez, R., Koo, E. H. & Selkoe, D. J. Interaction between amyloid precursor protein and presenilins in mammalian cells: implications for the pathogenesis of Alzheimer Disease. *Proc. Natl Acad. Sci. USA* **94**, 8208–8213 (1997).
- Ray, W. J. *et al.* Evidence for a physical interaction between presenilin and Notch. *Proc. Natl Acad. Sci. USA* **96**, 3263–3268 (1999).
- Thinakaran, G. *et al.* Stable association of the presenilin derivatives and absence of presenilin interactions with APP. *Neurobiol. Disease* **4**, 438–453 (1998).
- Capell, A. *et al.* The proteolytic fragments of the Alzheimer's Disease associated presenilin-1 form heterodimers and occur as a 100–150 kDa molecular mass complex. *J. Biol. Chem.* **273**, 3205–3211 (1998).
- Hay, J. C., Chao, D. S., Kuo, C. S. & Scheller, R. H. Protein interaction regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. *Cell* **89**, 149–158 (1997).
- Citron, M. *et al.* Additive effects of PS1 and APP mutations on secretion of the 42-residue amyloid beta-protein. *Neurobiol. Dis.* **5**, 107–116 (1998).
- Chen, F. Proteolytic derivative of Amyloid Precursor Protein accumulate in restricted and unpredicted intracellular compartments in the absence of functional presenilin 1 expression. *J. Biol. Chem.* (in the press).
- Zhang, L., Song, L. & Parker, E. M. Calpain inhibitor 1 increases beta-Amyloid peptide production by inhibiting the degradation of the substrate of gamma-secretase. *J. Biol. Chem.* **274**, 8966–8972 (1999).
- Howell, D. C. *Statistical Methods for Psychology* (Duxbury Press, California, 1992).

Supplementary information is available on *Nature's* World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

This work was supported by grants from the Medical Research Council of Canada, Alzheimer Association of Ontario, Howard Hughes Medical Research Foundation, Scottish Rite Charitable Foundation, Helen B. Hunter Fellowship (G.Y.), Peterborough Burgess Fellowship (E.A.R.), NIH (R.A. and L.F.); the National Institute of Aging Alzheimer Disease Center Grant (L.F.), University of Toronto Department of Medicine Postgraduate Fellowship (M.N.) and Japan Society for the Promotion of Science (T.K.). The PS2 D366A construct was from C. Haass. We thank R. Feldman and J.-F. Foncin for their work on the Nicastrin pedigrees, and C. Goutte and J. R. Preiss for sharing pre-publication data.

Correspondence and requests for materials should be addressed to P.St G.-H. (e-mail: p.physlop@utoronto.ca).

Non-quantized penetration of magnetic field in the vortex state of superconductors

A. K. Geim^{*†}, S. V. Dubonos^{*‡}, I. V. Grigorieva^{*†}, K. S. Novoselov^{*}, F. M. Peeters[§] & V. A. Schweigert^{§||}

^{*} University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

[†] Department of Physics, The University of Manchester, M13 9PL Manchester, UK

[‡] Institute for Microelectronics Technology, 142432 Chernogolovka, Russia

[§] Department of Physics, University of Antwerpen (UIA), B-2610 Antwerpen, Belgium

^{||} Institute of Theoretical and Applied Mechanics, 630090 Novosibirsk, Russia

As first pointed out by Bardeen and Ginzburg in the early sixties^{1,2}, the amount of magnetic flux carried by vortices in superconducting materials depends on their distance from the sample edge, and can be smaller than one flux quantum, $\Phi_0 = h/2e$ (where h is Planck's constant and e is the electronic charge). In bulk superconductors, this reduction of flux becomes negligible at sub-micrometre distances from the edge, but in thin films the effect may survive much farther into the material^{3,4}. But the effect has not been observed experimentally, and it is often assumed that magnetic field enters type II superconductors in units of Φ_0 . Here we measure the amount of flux introduced by individual vortices in a superconducting film, finding that the flux always differs substantially from Φ_0 . We have observed vortices that carry as little as $0.001\Phi_0$, as well as 'negative vortices', whose penetration leads to the expulsion of magnetic field. We distinguish two phenomena responsible for non-quantized flux penetration: the finite-size effect¹⁻⁴ and a nonlinear screening of the magnetic field due to the presence of a surface barrier. The latter effect has not been considered previously, but is likely to cause non-quantized penetration in most cases.

The magnetic properties of a superconductor, including its current-carrying capacity, are determined by the motion of flux through that superconductor as a whole; this motion involves propagation of flux not only through the bulk but also through the superconductor's edge. Because of the inevitable pinning in real superconductors, vortices can initially penetrate only at a finite (usually, mesoscopic) distance from the edge. This effectively creates an edge layer that serves as a reservoir of vortices that are subsequently injected further into the bulk, and there is growing evidence that such a layer significantly influences global superconducting properties^{5,6}. On the other hand, near-edge vortices are not exactly the same as vortices in the bulk because the distribution of electric currents around a vortex (that is, the vortex's structure) has to change owing to the presence of the edge¹⁻⁴.

One of the most directly observable consequences of the influence of an edge on a vortex is that its flux is no longer quantized and becomes smaller than Φ_0 (refs 1-4). This effect is particularly important in the case of thin films, where the screening is strongly suppressed and non-exponential^{3,4}. Although this flux reduction has been known theoretically for several decades, such vortices (carrying a fraction of Φ_0) have never been observed or inferred in an experiment. This provided the original motivation for our work, as we found a way to address the issue by making use of ballistic Hall magnetometry^{7,8}. This technique allows accurate magnetization measurements on micrometre-sized superconductors, where the edge effects can be dominant.

Figure 1 shows typical behaviour that we observed for the initial stages of field penetration in relatively large (15- μm), thin-film superconductors. Curve a shows magnetic flux penetrating inside a sample in a sequence of steps, such that each step corresponds to a

vortex or a number of vortices jumping inside; such behaviour is in agreement with general expectations. However, a more careful look reveals that the step height is not quantized, and that some jumps are smaller than Φ_0 . We postpone a discussion of this observation and now refer to another (nominally similar) sample in Fig. 1. Curve b (for this sample) reveals a completely different picture, which we have observed for many other samples. Here, after the initial region of the full Meissner effect, the flux enters the film relatively smoothly and, only after several flux jumps, the behaviour becomes qualitatively similar to the one shown in curve a. On the smooth part of curve b, the flux jumps correspond to a minor fraction of Φ_0 . Moreover, the first two jumps are negative, indicating that the superconductor expels magnetic field when a vortex jumps inside. The influence of the edge¹⁻⁴ discussed above can decrease the amplitude of flux jumps and is partly (see below) responsible for non-quantized steps. However, the existence of negative flux jumps is unexpected and seemingly makes no sense.

To understand the origin of the negative jumps as well as the reason why similar samples exhibit such different behaviour, we performed a number of experiments using various sample geometries. The results are summarized in Fig. 2, where we try to simplify the situation as much as possible by using relatively small disks and by examining only the penetration of the first vortex. The advantage of using such small samples is that bulk pinning becomes negligible compared to interaction of vortices with the edge and, as a result, the first vortex comes right to the disk's centre^{8,9}. Therefore, we can study the penetration of an individual vortex at the same, well defined, distance ($R = D/2$) from the edge. As seen from Fig. 2, the amount of flux carried by vortices entering the disk depends on the roughness of the edge of the disk. The disk shown in Fig. 2a, with a smooth edge, exhibits a negative flux jump when the first vortex

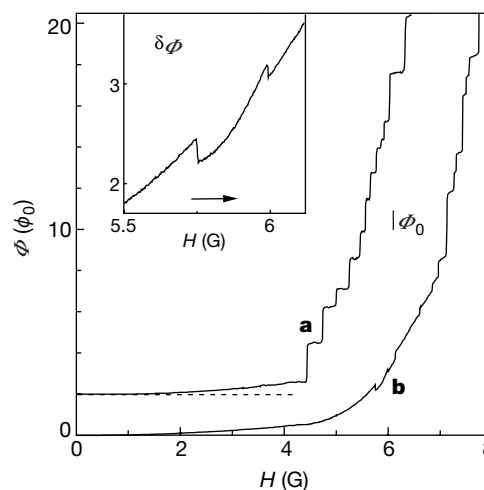


Figure 1 Penetration of perpendicular magnetic field in a thin superconducting film. The curves show the amount of flux Φ , as a function of increasing field H , inside two aluminium disks of diameter $D \approx 15 \mu\text{m}$ and thickness $h \approx 0.1 \mu\text{m}$ at $T \approx 0.5 \text{ K}$. Due to bulk pinning, which is rather weak⁹ but still present, entering vortices jump no farther than a few micrometres from the edge (we observe hysteresis due to bulk pinning if $D > 4 \mu\text{m}$). This makes measurements for the larger disks essentially equivalent to a study of flux penetration in a πD -long strip of an identical macroscopic film. Initially, the samples were cooled in zero field. Special care was taken to avoid 'freezing-in' any vortices; the absence of such vortices was verified by observing a symmetric response for the opposite field direction. The measurements were performed using ballistic Hall magnetometry (Fig. 2). For convenience, we define Φ so that it has zero slope in the low-field limit where $M \propto H$ (such a notation ignores the amount of flux in the -1a layer for the ideal Meissner state; taking the latter flux into account would only lead to an additional, constant slope for $\Phi-H$ curves). The absolute scale along the Φ axis is determined with an experimental accuracy of about 10%. Curve a is shifted for clarity. Inset, magnified view of part of curve b, exhibiting negative flux jumps.

enters it. On the other hand, if we introduce a sharp defect at the edge of the disk (as in the disk shown in Fig. 2c), the more-or-less expected behaviour is recovered. Figure 2b shows an intermediate case, where a vortex brings in practically no flux. We note that the leaving vortices carry away approximately the same amount of flux (about $0.7\phi_0$), independent of edge roughness. If no special care is taken, our disks usually have a few tiny cuts at the edge⁷⁻⁹ that occur because of ripping of the evaporated film during the standard lift-off procedure. To avoid this, the disks in Fig. 2 were fabricated using a double-layer resist. A retrospective analysis has shown that the sample that produced curve a in Fig. 1 has a jagged edge, while the edge of the sample that produced curve b is smooth.

Figure 2 shows clearly that the negative jumps are somehow connected with the stronger bending of the flux-field (Φ - H) curves and with the increase of the penetration field H_p for the first vortex. The smoother the edge, the longer the Meissner state persists and the more strongly the Meissner curve bends upwards. The bending means that the magnetic field penetrating in the near-edge layer cannot be described by the London model, which requires $M \propto H$ and $\Phi(H) = 0$ (see Fig. 1 legend). The nonlinear Meissner effect observed for samples with smooth edges is due to the presence of a high surface barrier that allows superconductors to persist in a metastable (superheated) state⁸⁻¹⁰. Figure 2 shows that the edge roughness strongly suppresses the barrier for vortex entry, while the exit barrier is influenced relatively weakly.

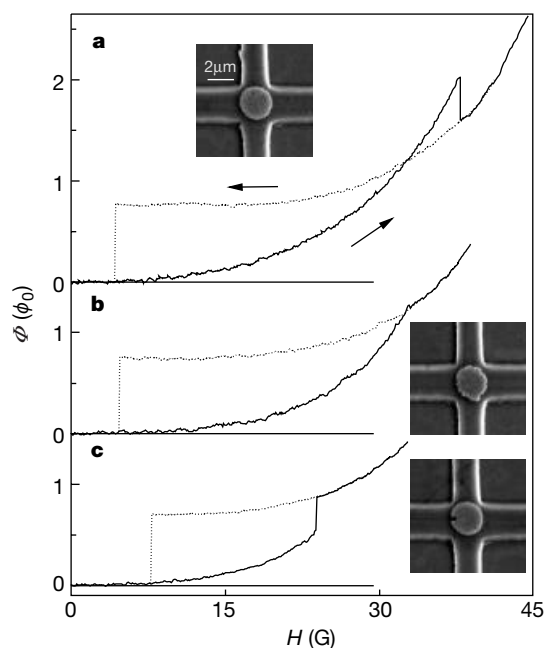


Figure 2 Penetration and expulsion of the first vortex in superconductors with various edge roughnesses. Solid and dotted curves were obtained by sweeping the magnetic field up and down, respectively. The hysteresis is due to a surface barrier and, until a vortex enters or leaves the disks, the curves are reproducible for sweeps in both directions. Insets, micrographs of the studied Al disks ($D \approx 2 \mu\text{m}$ and $h \approx 0.15 \mu\text{m}$; $T \approx 0.5 \text{ K}$). The rough edges for the disks shown in **b** and **c** were intentionally drawn by electron-beam lithography. The disks are placed on top of Hall crosses (see the micrographs) made from a high-mobility two-dimensional electron gas. Such probes measure the average magnetic field in the central area of the Hall cross⁷. The Hall magnetometers can be considered as flux meters with a square detection loop, in the centre of which a superconducting sample is placed. The superconducting coherence length ξ ($T = 0 \text{ K}$) for the disks' material is about $0.25 \mu\text{m}$ and the magnetic penetration length $(0\lambda \approx 70 \text{ nm})$, that is, the material is a type I superconductor (Ginzburg-Landau parameter, $\kappa = \lambda/\xi \approx 0.3$). We note that thin films in a perpendicular magnetic field behave more like type II superconductors, and exhibit vortex structures⁸⁻¹¹. We can move into the true type II regime by using less-pure Al, and have observed the discussed behaviour also for $\kappa > 1$. However, the reduced screening due to unavoidably larger λ for large κ led to rapid deterioration of our experimental resolution, and here we present the data for lower κ .

In order to explain the origin of negative flux jumps, we will refer to the theoretical curves in Fig. 3a that show the magnetization response for a disk with parameters close to those in Fig. 2a. The important feature to notice is the intersection of magnetization curves for different vortex states. The curve crossings occur in a metastable regime and mean that, as the field increases, the preceding vortex state can accommodate more flux (that is, can be less diamagnetic) than the following, more energetically favourable, state with a larger L . This leads to the possibility of vortex jumps that can bring inside a superconductor any amount of flux from zero to about $\pm \phi_0$, depending on H_p . The theory allows metastable states, but is not able to predict at what value of H they become unstable and a vortex jumps in or out^{10,11}. In our experiments, only a part of the full length of the theoretical curves is realized. Indeed, we have never seen the intersection of magnetization curves with decreasing H , which could lead to an alternative situation where the flux inside a superconductor increases when a vortex leaves.

To elucidate the physical processes behind the observed negative jumps, we consider the distribution of magnetic field before and after such jumps. In Fig. 3b we plot the field distribution for the superheated Meissner state and the state with one vortex inside at the same applied field. For clarity, we will discuss the case of a superconducting cylinder rather than a disk. The former geometry keeps the essential physical processes intact, but allows us to avoid obscuring demagnetization effects. This also serves to illustrate the fact that negative flux jumps are not exclusive to the thin-film geometry. In the superheated state, the external field strongly suppresses the order parameter near the edge, so that the field penetrates inside the cylinder by as much as a few λ deeper than the

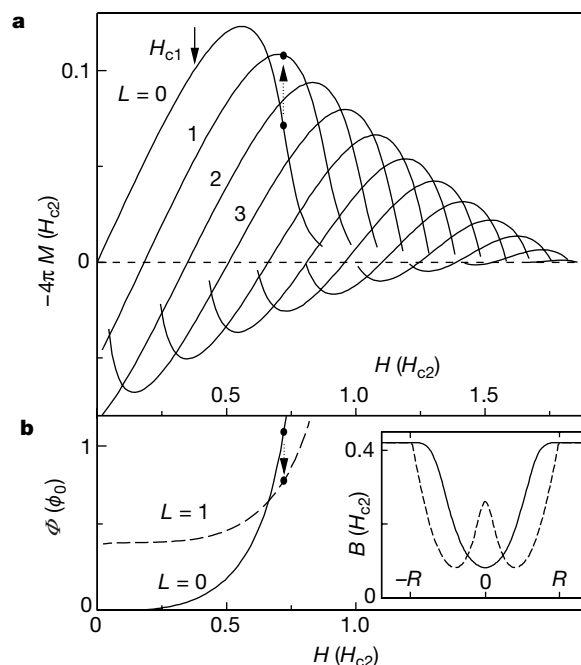


Figure 3 Magnetic response of mesoscopic superconductors found theoretically. **a**, Calculated magnetization response for a disk with $D = 8\xi$, $h = 0.1\xi$ and $\lambda \approx 0.28\xi$. The curves are found by solving numerically the full three-dimensional set of Ginzburg-Landau equations¹⁰. L is the fluxoid number which, for simplicity, can be considered as the number of vortices inside the disk. **b**, The amount of flux Φ inside the disk replotted from the M - H dependence. The solid arrow marks the low critical field H_{c1} where the Meissner state ($L = 0$) and the state with one vortex ($L = 1$) have the same free energy. Dots on the curves show an example of the situation where vortex entry could lead to flux expulsion. The inset shows the radial distribution of magnetic field B inside a superconducting cylinder ($D = 20\xi$ and $\lambda \approx 2\xi$) for a similar situation, in which an entering vortex reduces the amount of flux inside the cylinder. The Meissner state and the vortex state are shown by the solid and the dashed curve, respectively.

London model allows. The higher the applied field, the more strongly the order parameter is suppressed and the more flux is accumulated in the near-edge layer. On the other hand, when the vortex jumps inside, the screening is restored to a significant extent and, accordingly, there is less flux in the near-edge layer (Fig. 3). The competition between the flux expelled from this layer and the vortex's flux determines the sign and amplitude of flux jumps. Despite the deceptive simplicity of this explanation, there is no simple way to explain why the vortex entry restores screening while the field at the edge remains the same. This is a nonlinear property of superconductors.

We now turn to the question of why the flux jumps are not quantized, even when the surface barrier is suppressed by edge roughness, and why the distance between the curves with and without a vortex is less than ϕ_0 (see Fig. 2). The latter implies that the vortex's flux even in equilibrium (that is, not only the corresponding flux jumps) is considerably less than ϕ_0 . This observation can be explained by the changes in the structure of near-edge vortices predicted in refs 1–4. Figure 4 plots the measured amount of flux ϕ carried by a vortex versus its distance from the disk's edge. We can see that all our data for different samples and temperatures fall on a single curve, if plotted in units of the effective penetration length, $\lambda_{\text{eff}} = \lambda / \sqrt{1 + \alpha^2}$. There is also excellent agreement with the corresponding theoretical dependence. We note that, for a typical experimental situation, h and λ are about $0.1 \mu\text{m}$, and it is very unlikely that a vortex can jump farther than $1 \mu\text{m}$ from the edge before being stopped by pinning, even in samples with low pinning. According to Fig. 4, in such a case the flux carried by vortices is reduced to about $0.5\phi_0$. Only vortices located as far as $100 \mu\text{m}$ away from the film edge have their flux quantized with an accuracy better than 1%.

We have shown that there are two independent effects that lead to non-quantized penetration of magnetic field in type II superconductors. The first (theoretically established a long time ago, but

never observed and often perceived as small) arises due to changes in the structure of near-edge vortices. This should be important in thin films and, in our opinion, may account for a number of unexplained observations. The second, unexpected, effect is more general, and appears owing to the inevitable presence of barriers for flux motion through a superconducting boundary (for example, Bean–Livingston barriers). If such barriers are sufficiently high, nonlinear screening can lead to the extreme situation, causing ‘negative vortices’; but if this is not the case, surface barriers can still prevent the quantized penetration. One or both of the above effects can be expected in many—if not most—relevant experimental situations. □

Received 16 June; accepted 12 July 2000.

1. Bardeen, J. Quantization of flux in a superconducting cylinder. *Phys. Rev. Lett.* **7**, 162–163 (1961).
2. Ginzburg, V. L. Magnetic flux quantization in a superconducting cylinder. *Sov. Phys. JETP* **15**, 207–209 (1962).
3. Fetter, A. L. Flux penetration in a thin superconducting disk. *Phys. Rev. B* **22**, 1200–1213 (1980).
4. Kogan, V. G. Pearl's vortex near the film edge. *Phys. Rev. B* **49**, 15874–15878 (1994).
5. White, W. R., Kapitulnik, A. & Beasley, M. R. Collective vortex motion in Mo–Ge superconducting thin films. *Phys. Rev. Lett.* **70**, 670–673 (1993).
6. Fuchs, D. T. *et al.* Transport properties governed by surface barriers in BiSrCaCuO. *Nature* **391**, 373–376 (1998).
7. Geim, A. K. *et al.* Ballistic Hall micromagnetometry. *Appl. Phys. Lett.* **71**, 2379–2381 (1997).
8. Geim, A. K., Dubonos, S. V., Lok, J. G. S., Henini, M. & Maan, J. C. Paramagnetic Meissner effect in small superconductors. *Nature* **396**, 144–146 (1998).
9. Geim, A. K. *et al.* Fine structure in magnetization of individual fluxoid states. *Phys. Rev. Lett.* **85**, 1528–1531 (2000).
10. Schweigert, V. A. & Peeters, P. M. Flux penetration and expulsion in thin superconducting disks. *Phys. Rev. Lett.* **83**, 2049–2052 (1999).
11. Palacios, J. J. Metastability and paramagnetism in superconducting disks. *Phys. Rev. Lett.* **84**, 1796–1799 (2000).

Acknowledgements

We thank V. Kogan for discussions, and FOM, INTAS and NATO for financial support.

Correspondence and requests for materials should be addressed to A.K.G. (e-mail: geim@sci.kun.nl).

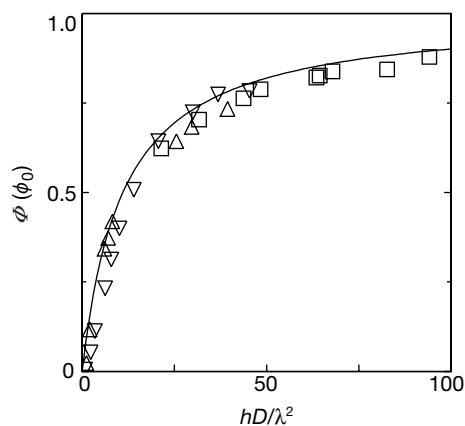


Figure 4 The amount of magnetic flux associated with a vortex in the centre of a thin superconducting disk in equilibrium. The solid curve is an approximate theoretical dependence, $\phi \approx \phi_0 \gamma / (\alpha + \gamma)$, found numerically in the limit $h \ll \lambda$, where $\gamma = hD/\lambda^2$ and $\alpha \approx 11.0$ (ref. 3, and V.A.S. *et al.*, manuscript in preparation). Different symbols show experimental data for three disks with D (in μm) ~ 2 (upright triangles), ~ 2.4 (inverted triangles) and ~ 4 (squares), h from about 0.13 to $0.17 \mu\text{m}$ and the superconducting parameters as in Fig. 2. Because of the surface barrier that is always present (even for a rough edge), we cannot directly determine H_{c1} and, therefore, the amount of flux carried by a vortex in equilibrium. To this end, we notice that the theoretical curves (M – H and Φ – H) for $L = 0$ and 1 are nearly parallel below H_{c1} (Fig. 3) and, hence, the amount of flux associated with vortex exit is sufficiently close—within our experimental uncertainty of 10%—to the vortex's flux in equilibrium. So we have measured the amplitude of flux jumps for the vortex exit. To obtain different data points for each of the disks, we varied the penetration length (λ) by changing the temperature from 0.4 K to close to $T_c \approx 1.25 \text{ K}$. No fitting parameters were used, except for a slight adjustment ($\leq 10\%$) of the absolute scale along the ϕ axis for each of the disks.

Nanomechanical oscillations in a single- C_{60} transistor

Hongkun Park^{*‡}, Jiwoong Park[†], Andrew K. L. Lim^{*}, Erik H. Anderson[‡], A. Paul Alivisatos^{*‡} & Paul L. McEuen^{†‡}

^{*} Department of Chemistry and [†] Department of Physics, University of California at Berkeley and [‡] Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

The motion of electrons through quantum dots is strongly modified by single-electron charging and the quantization of energy levels^{1,2}. Much effort has been directed towards extending studies of electron transport to chemical nanostructures, including molecules^{3–8}, nanocrystals^{9–13} and nanotubes^{14–17}. Here we report the fabrication of single-molecule transistors based on individual C_{60} molecules connected to gold electrodes. We perform transport measurements that provide evidence for a coupling between the centre-of-mass motion of the C_{60} molecules and single-electron hopping¹⁸—a conduction mechanism that has not been observed previously in quantum dot studies. The coupling is manifest as quantized nano-mechanical oscillations of the C_{60} molecule against the gold surface, with a frequency of about 1.2 THz . This value is in good agreement with a simple theoretical

[§] Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

estimate based on van der Waals and electrostatic interactions between C_{60} molecules and gold electrodes.

Single- C_{60} transistors were prepared by depositing a dilute toluene solution of C_{60} onto a pair of connected gold electrodes fabricated using electron-beam lithography. A break-junction technique¹¹ was then used to create a gap between these electrodes by the process of electromigration. The typical lateral size of the fabricated electrodes was of the order of 100 nm at the point of the gap formation, and the height of the electrodes was 15 nm. Scanning electron microscope images of fabricated electrodes reveal that the gap between two electrodes is not uniform and that the narrowest gap is formed only between small protrusions (≤ 10 nm) of two gold electrodes. Current–voltage measurements¹¹ of these electrodes at cryogenic temperatures without deposited C_{60} molecules show that the size of the gap is consistently about 1 nm. In a significant fraction of the C_{60} devices, the conductance of the junction after initial breaking is substantially enhanced compared to devices with no C_{60} deposited, indicating that C_{60} molecules reside in the junction. The entire structure was defined on a SiO_2 insulating layer on top of a degenerately doped silicon wafer that serves as a gate electrode that modulates the electrostatic potential of C_{60} . A schematic diagram of an idealized single- C_{60} transistor is shown in the lower inset of Fig. 1.

Figure 1 presents representative current–voltage (I – V) curves obtained from a single- C_{60} transistor at different gate voltages (V_g). The device exhibited strongly suppressed conductance near zero bias voltage followed by step-like current jumps at higher voltages. The voltage width of the zero-conductance region (conductance gap) could be changed in a reversible manner by changing V_g . In ten devices prepared from separate fabrication runs, the conductance gap could be reduced to zero by adjusting V_g , although the gate voltage at which the conductance gap closed (V_c) varied from device to device.

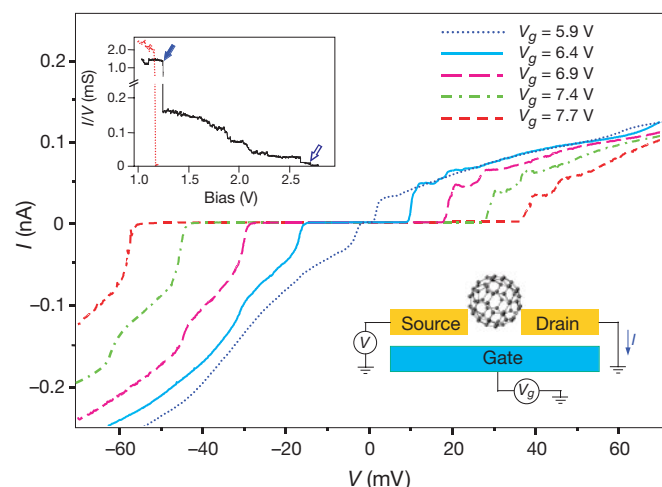


Figure 1 Current–voltage (I – V) curves obtained from a single- C_{60} transistor at $T = 1.5$ K. Five I – V curves taken at different gate voltages (V_g) are shown. Single- C_{60} transistors were prepared by first depositing a dilute toluene solution of C_{60} onto a pair of connected gold electrodes. A gap of ~ 1 nm was then created using electromigration-induced breaking of the electrodes. Upper inset, a large bias was applied between the electrodes while the current through the connected electrode was monitored (black solid curve). After the initial rapid decrease (solid arrow), the conductance stayed above ~ 0.05 mS up to ~ 2.0 V. This behaviour was observed in most single- C_{60} transistors, but it was not observed when no C_{60} solution was deposited (red dotted curve). The bias voltage was increased until the conductance fell low enough to ensure that the current through the junction was in the tunnelling regime (open arrow). The low bias measurements shown in the main panel were taken after the breaking procedure. Lower inset, an idealized diagram of a single C_{60} -transistor formed by this method.

Figures 2 and 3 show two-dimensional plots of differential conductance ($\partial I/\partial V$) as a function of both V and V_g for four different devices. Peaks in $\partial I/\partial V$, which correspond to the step-like features in Fig. 1, show up as lines in these plots. As seen clearly in Figs 2 and 3, the size of the conductance gap and the $\partial I/\partial V$ peak positions evolve smoothly as V_g is varied. As the gate voltage was varied farther away from V_c in both positive and negative directions, the conductance gap continued to widen and the maximum observed gap exceeded 270 mV. Many $\partial I/\partial V$ peaks outside the conductance gap are also observed.

The V_g -dependent features described above were not observed in devices when C_{60} was not deposited on the electrodes. In addition, the coverage of C_{60} on the electrodes was such that only about 10% of more than 300 fabricated electrodes show I – V characteristics that are different from a simple tunnel junction without C_{60} . This low C_{60} coverage ensures that the probability of finding multiple C_{60} molecules bridging two electrodes is small. Furthermore, many different devices exhibited similar conductance characteristics that are consistent with a single nanometre-sized object bridging two electrodes, as explained in detail below¹. Although C_{60} could not be imaged directly in these devices because of its small size (~ 7 Å in diameter), these experimental observations indicate that individual C_{60} molecules are responsible for the conductance features observed in the experiment.

The global patterns observed in Figs 1–3 can be understood using ideas borrowed from the Coulomb blockade model developed for the analysis of quantum-dot transport¹. The conductance gap observed in the data is a consequence of the finite energy required to add (remove) an electron to (from) C_{60} . This energy cost arises from the combined effect of single-electron charging of C_{60} and the quantized excitation spectrum of the C_{60} -transistor system. The maximum observed gap in the experiment indicates that the charging energy of the C_{60} molecule in this geometry can exceed 270 meV.

The conductance gap changes reversibly as a function of V_g because a more positive gate voltage stabilizes an additional electron

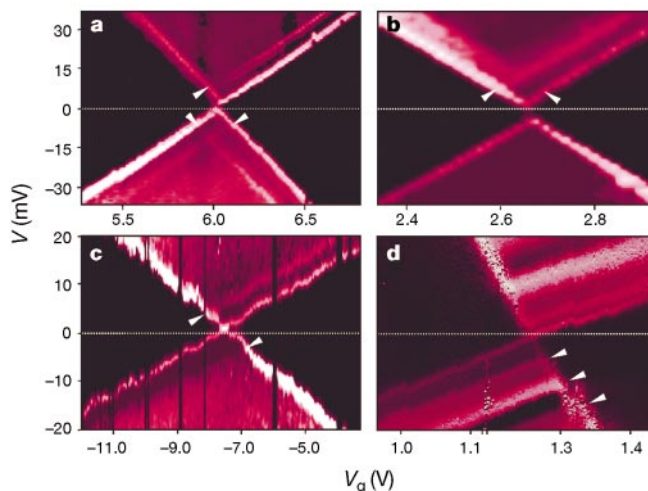


Figure 2 Two-dimensional differential conductance ($\partial I/\partial V$) plots as a function of the bias voltage (V) and the gate voltage (V_g). Data were obtained from four different devices prepared from separate fabrication runs. The dark triangular regions correspond to the conductance gap, and the bright lines represent peaks in the differential conductance. **a–d**, The differential conductance values are represented by the colour scale, which changes from black (0 nS) through pink to white (white representing 30 nS in **a**, **b** and **c** and 5 nS in **d**). The white arrows mark the point where $\partial I/\partial V$ lines intercept the conductance gap. During the acquisition of data in **d**, one 'switch' where the entire $\partial I/\partial V$ characteristics shift along the V_g axis occurred at $V_g = 1.15$ V. The right portion of the plot **d** is shifted along the V_g axis to preserve the continuity of the lines.

on C_{60} . The conductance gap disappears at $V_g = V_c$ where the total energy of the system is the same for two different C_{60} charge states. When the gate voltage traverses V_c in the positive direction, the equilibrium number of charges on C_{60} changes by one electron from C_{60}^{n-} to $C_{60}^{(n+1)-}$, where n designates the number of charges on C_{60} . It is determined by both V_g and the local electrochemical environment, that is, the work function of the metal electrode and the local charge distribution around C_{60} . While the value of n cannot be determined solely from our experimental data, previous electrochemical and photoelectron spectroscopic studies of C_{60} on gold suggest that n is most likely to be zero or one¹⁹.

The position of each $\partial I/\partial V$ peak outside the conductance gap in Figs 2 and 3 provides detailed information on the quantized excitations of the single- C_{60} transistor system¹. These $\partial I/\partial V$ peaks appear when a new quantized excitation becomes energetically accessible, providing an electron-tunnelling pathway between C_{60} and the gold electrodes. Specifically, each $\partial I/\partial V$ peak on the $V_g < V_c$ side signifies an opening of a new tunnelling pathway where an electron hops onto C_{60}^{n-} to generate $C_{60}^{(n+1)-}$ in its ground or excited states; these peaks therefore probe the excitation energies of the $C_{60}^{(n+1)-}$ ion. Each $\partial I/\partial V$ peak that appears at $V_g > V_c$ occurs when an electron hops off $C_{60}^{(n+1)-}$ to generate C_{60}^{n-} ; these peaks thus probe the ground and excited states of C_{60}^{n-} . The energy of these quantized excitations can be determined from the bias voltage at which they intercept the conductance gap¹, as shown by the white arrows in Fig. 2.

A remarkable common feature of the different devices is that a quantized excitation is universally observed with an energy of approximately 5 meV. Moreover, this excitation is observed on both sides of V_c in most devices, indicating that it is an excitation of both charge states of C_{60} . The exact value of this energy quantum varied from device to device and ranged from 3 to 7 meV. In some devices, multiple $\partial I/\partial V$ features with almost identical spacing appear, as seen in Fig. 2d.

The observed 5-meV excitation could arise from many possible degrees of freedom of the single- C_{60} transistor system. One possibility, which has often been invoked in other nanometre-scale systems, is the excited electronic states of the system. However, this possibility is highly unlikely here because the 5-meV excitation is the same for both charge states of C_{60} and also because multiple excitations with the same spacing are observed. Although the exact electronic-level structures of the C_{60}^{n-} ions are not known experimentally, theoretical calculations predict that the electronic states do not follow such behaviour^{19,20}.

A more natural candidate is a vibrational excitation of the C_{60}

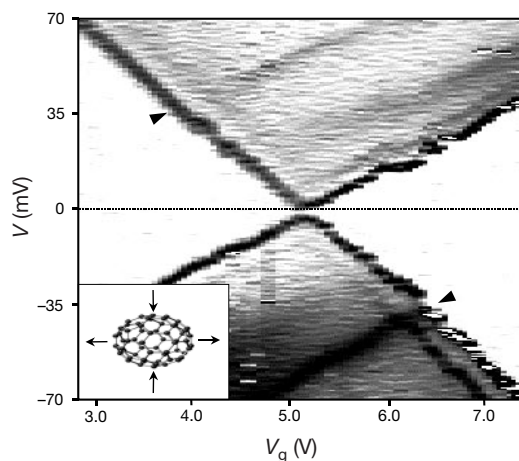


Figure 3 A differential conductance plot showing a larger bias-voltage range than those in Fig. 2. Here two $\partial I/\partial V$ lines that intercept the conductance gap at $V = 35$ mV are seen clearly (arrows). The energy quantum of this excitation closely matches that of the C_{60} internal vibrational mode shown in the inset.

system coupled to an electron tunnelling on and off C_{60} . The observation of multiple $\partial I/\partial V$ features with identical spacing would then result from the excitation of integer numbers of vibrational quanta. Moreover, these vibrational modes would be present irrespective of the charge state of the C_{60} .

The internal vibrational modes of the free C_{60} molecule have been extensively studied, both theoretically and experimentally^{19,21}. The lowest-energy mode is one with a vibrational quantum of 33 meV and corresponds to the C_{60} deformation from a sphere to a prolate ellipsoid, as shown in the inset to Fig. 3. In Fig. 3, an excitation that probably corresponds to this mode can indeed be seen with an energy of about 35 meV. However, internal vibrational modes cannot account for the observed 5-meV features.

Another possibility is the centre-of-mass oscillation of C_{60} within the confinement potential that binds it to the gold surface, as shown in Fig. 4. To our knowledge, this mode has not been directly measured experimentally. However, previous theoretical and experimental studies have shown that C_{60} is held tightly on gold by van der Waals interactions, with a C_{60} –gold binding energy of about 1 eV and a distance of about 6.2 Å between the C_{60} centre and the gold surface^{19,22,23}. Assuming that the C_{60} –gold interaction potential can be expressed by the Lennard–Jones form²³, the above parameters can be used to determine the shape of the potential that describes the C_{60} –gold binding. This calculation indicates that the C_{60} –gold binding near the equilibrium position can be approximated very well by a harmonic potential with an estimated force constant of $k \approx 70$ N m⁻¹, as is shown schematically in Fig. 4. This force constant and the mass M of the C_{60} molecule yield a vibrational frequency of $f = 1/2\pi(k/M)^{1/2} \approx 1.2$ THz and a vibrational quantum of $hf \approx 5$ meV, where h is the Planck constant.

Adding an additional electron to C_{60} compresses the C_{60} –surface bond due to the interaction between the C_{60} ion and its image charge in the metal. A simple estimate indicates that an additional electron on C_{60} results in the shortening of the C_{60} –surface distance by $\delta \approx 4$ pm, but it does not significantly change the vibrational

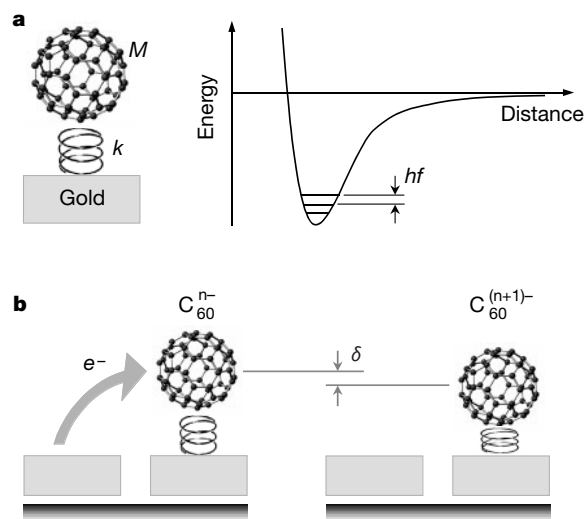


Figure 4 Diagram of the centre-of-mass oscillation of C_{60} . **a**, A C_{60} molecule is bound to the gold surface by the van der Waals and electrostatic interaction. The interaction potential is shown schematically alongside. The potential near the equilibrium position can be approximated well by a harmonic potential with a force constant k . This harmonic potential gives quantized energy levels with frequency $f = 1/2\pi(k/M)^{1/2}$. Here M represents the mass of C_{60} and h is the Planck constant. **b**, When an electron jumps on to C_{60}^{n-} , the attractive interaction between the additional electron and its image charge on gold pulls the C_{60} ion closer to the gold surface by the distance δ . This electrostatic interaction results in the mechanical motion of C_{60} .

frequency. By comparison, the r.m.s. (root mean square) displacement x_m of the C_{60} molecule in the m th vibrational level is given by $x_m = (2m + 1)^{1/2} x_0$, where $x_0 = (\hbar f/k)^{1/2} \approx 3$ pm. Although the exact values of these simple estimates change when the second metal electrode is included in the model, the qualitative conclusions of the model remain essentially the same. The present estimates pertain to the situation where the coupling between C_{60} and two electrodes is strongly asymmetric.

The C_{60} -surface vibration discussed above can account for the 5-meV conductance features in a unifying fashion. The first $\partial I/\partial V$ peak at the boundary of the conductance-gap region is observed when an electron hops on or off C_{60} with the system staying in the ground vibrational level. Additional $\partial I/\partial V$ peaks on the $V_g < V_c$ side appear when an electron hops onto C_{60}^{n-} to generate $C_{60}^{(n+1)-}$ in excited vibrational states. The $\partial I/\partial V$ peaks on the $V_g > V_c$ side signify, on the other hand, an event where an electron hops off $C_{60}^{(n+1)-}$, leaving C_{60}^{n-} in excited vibrational levels. Multiple $\partial I/\partial V$ peaks on the same side of V_c indicate that multiple vibrational quanta are excited.

This process is thus reminiscent of the Franck-Condon processes encountered in electron-transfer and light-absorption processes in molecules, where the vibrational excitation accompanies the electronic motion²⁴. Within the harmonic approximation, the vibrational matrix elements for these processes can readily be calculated, and the ratio δ/x_0 determines the number of vibrational quanta typically excited by the tunnelling electron. According to the estimates discussed above, $\delta/x_0 \approx 1$ in a single- C_{60} transistor. The number of $\partial I/\partial V$ peaks visible in Fig. 2 in general confirms this expectation as only a few $\partial I/\partial V$ peaks are observed in most devices.

One device that does not follow this general trend is the one shown in Fig. 2d. As described previously, this device exhibits many $\partial I/\partial V$ peaks on both sides of V_c . In addition, the peak intensities do not show the simple variations expected from the single-mode Franck-Condon situation²⁴. The anomalous behaviour may be related to the highly asymmetric coupling of C_{60} and the two electrodes in this particular device. This asymmetry is demonstrated by the different slopes of the upward and downward $\partial I/\partial V$ lines in the $V-V_g$ plane. The variations of peak intensities may be due to the presence of other degrees of freedom in the system, such as the C_{60} motion perpendicular to the surface normal.

Unexplained features exist in other devices as well. In the data in Fig. 2a, a small (≤ 1 meV) energy splitting is observed for many of the lines. This splitting may arise from the C_{60} centre-of-mass motion perpendicular to the surface normal discussed above. Unfortunately, the nature of the potential for this motion is not known, owing to the lack of detailed knowledge of the electrode geometry near C_{60} , and quantitative support of this assignment is thus lacking at present.

The transport measurements presented here demonstrate that single-electron-tunnelling events can be used both to excite and probe the motion of a molecule: indeed, the single- C_{60} transistor behaves as a high-frequency nanomechanical oscillator. Furthermore, the oscillations of the C_{60} molecule must be treated in a quantized fashion, showing that this is a true quantum 'mechanical' system. We expect that the coupling between the quantized electronic and mechanical degrees of freedom will be generically important in electron transport through nanomolecular systems^{25,26}. □

Received 17 March; accepted 30 June 2000.

- Grabert, H. & Devoret, M. H. *Single Charge Tunneling* (Plenum, New York, 1992).
- Sohn, L. L., Kouwenhoven, L. P. & Schön, G. *Mesoscopic Electron Transport* (Kluwer Academic, Dordrecht, 1997).
- Bumm, L. A. *et al.* Are single molecular wires conducting? *Science* **271**, 1705–1707 (1996).
- Reed, M. A., Zhou, C., Muller, C. J., Burgin, T. P. & Tour, J. M. Conductance of a molecular junction. *Science* **278**, 252–254 (1997).
- Datta, S. *et al.* Current-voltage characteristics of self-assembled monolayers by scanning tunneling microscopy. *Phys. Rev. Lett.* **79**, 2530–2533 (1997).
- Porath, D., Levi, Y., Tarabiah, M. & Millo, O. Tunneling spectroscopy of isolated C_{60} molecules in the presence of charging effects. *Phys. Rev. B* **56**, 9829–9833 (1997).

- Joachim, C. & Gimzewski, J. K. An electromechanical amplifier using a single molecule. *Chem. Phys. Lett.* **265**, 353–357 (1997).
- Stipe, B. C., Rezaei, M. A. & Ho, W. Single-molecule vibrational spectroscopy and microscopy. *Science* **280**, 1732–1735 (1998).
- Klein, D. L., McEuen, P. L., Bowen Katari, J. E., Roth, R. & Alivisatos, A. P. An approach to electrical studies of single nanocrystals. *Appl. Phys. Lett.* **68**, 2574–2576 (1996).
- Klein, D. L., Roth, R., Lim, A. K. L., Alivisatos, A. P. & McEuen, P. L. A single-electron transistor made from a cadmium selenide nanocrystal. *Nature* **389**, 699–701 (1997).
- Park, H., Lim, A. K. L., Park, J., Alivisatos, A. P. & McEuen, P. L. Fabrication of metallic electrodes with nanometer separation by electromigration. *Appl. Phys. Lett.* **75**, 301–303 (1999).
- Banin, U., Cao, Y., Katz, D. & Millo, O. Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400**, 542–544 (1999).
- Kim, S.-H., Medeiros-Ribeiro, G., Ohlberg, D. A. A., Williams, R. S. & Heath, J. R. Individual and collective electronic properties of Ag nanocrystals. *J. Phys. Chem. B* **103**, 10341–10347 (1999).
- Bockrath, M. *et al.* Single-electron transport in ropes of carbon nanotubes. *Science* **275**, 1922–1925 (1997).
- Tans, S. J. *et al.* Individual single-wall carbon nanotubes as quantum wires. *Nature* **386**, 474–477 (1997).
- Tans, S. J., Verschuere, A. R. M. & Dekker, C. Room-temperature transistor based on a single carbon nanotube. *Nature* **393**, 49–52 (1998).
- Bockrath, M. *et al.* Luttinger-liquid behaviour in carbon nanotubes. *Nature* **397**, 598–601 (1999).
- Gorelik, L. Y. *et al.* Shuttle mechanism for charge transfer in Coulomb blockade nanostructures. *Phys. Rev. Lett.* **80**, 4526–4529 (1998).
- Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, New York, 1996).
- Green, W. H. Jr *et al.* Electronic structures and geometries of C_{60} anions via density functional calculations. *J. Phys. Chem.* **100**, 14892–14898 (1996).
- Heid, R., Pintschovius, L. & Godard, J. M. Eigenvectors of internal vibrations of C_{60} : Theory and experiment. *Phys. Rev. B* **56**, 5925–5936 (1997).
- Chavy, C., Joachim, C. & Altibelli, A. Interpretation of STM images: C_{60} on the gold (110) surface. *Chem. Phys. Lett.* **214**, 569–575 (1993).
- Ruoff, R. S. & Hickman, A. P. Van der Waals binding of fullerenes to a graphite plane. *J. Phys. Chem.* **97**, 2494–2496 (1993).
- Schatz, G. C. & Ratner, M. A. *Quantum Mechanics in Chemistry* (Prentice Hall, Englewood Cliffs, 1993).
- Poncharal, P., Wang, Z. L., Ugarte, D. & De Heer, W. A. Electrostatic deflections and electromechanical resonances of carbon nanotubes. *Science* **283**, 1513–1516 (1999).
- Reulet, B. *et al.* Bolometric detection of mechanical bending waves in suspended carbon nanotubes. Preprint cond-mat/9907486 at (<http://xxx.lanl.gov>) (1999).

Acknowledgements

We thank M. S. Fuhrer and N. S. Wingreen for discussions and advice. This work was supported by the US Department of Energy. E.A. was also partially supported by DARPA. Correspondence and requests for materials should be addressed to P.M. (e-mail: mceuen@socrates.berkeley.edu).

Gas-phase production and photoelectron spectroscopy of the smallest fullerene, C_{20}

Horst Prinzbach*, Andreas Weiler*, Peter Landenberger*, Fabian Wahl*, Jürgen Wörth*, Lawrence T. Scott†, Marc Gelmont†, Daniela Olevano‡ & Bernd v. Issendorff ‡

* Institut für Organische Chemie und Biochemie, Albert-Ludwigs-Universität, 79104 Freiburg, Germany

† Department of Chemistry, Boston College, Merkert Chemistry Center, Chestnut Hill, Massachusetts 02467-3860, USA

‡ Fakultät für Physik, Albert-Ludwigs-Universität, 79104 Freiburg, Germany

Fullerenes are graphitic cage structures incorporating exactly twelve pentagons¹. The smallest possible fullerene is thus C_{20} , which consists solely of pentagons. But the extreme curvature and reactivity of this structure have led to doubts about its existence and stability. Although theoretical calculations have identified, besides this cage, a bowl and a monocyclic ring isomer as low-energy members of the C_{20} cluster family², only ring isomers of C_{20} have been observed^{3–6} so far. Here we show that the cage-structured fullerene C_{20} can be produced from its perhydrogenated form (dodecahedrane $C_{20}H_{20}$) by replacing the hydrogen atoms

with relatively weakly bound bromine atoms, followed by gas-phase debromination. For comparison we have also produced the bowl isomer of C_{20} using the same procedure. We characterize the generated C_{20} clusters using mass-selective anion photoelectron spectroscopy; the observed electron affinities and vibrational structures of these two C_{20} isomers differ significantly from each other, as well as from those of the known monocyclic isomer. We expect that these unique C_{20} species will serve as a benchmark test for further theoretical studies.

The fullerene C_{20} , which may be regarded as the carbon transfiguration of "Plato's universe"⁷, is shown in Fig. 1 (compound 1). Unlike C_{60} , the fullerene C_{20} is not formed spontaneously in carbon condensation or cluster annealing processes, as has been demonstrated by extensive ion mobility measurements^{6,8}. It seemed to be a safe prediction that cage 1, if it can be produced at all, would have to be made from a precursor that bore a close structural resemblance to 1. Furthermore, it would be necessary to produce 1 in the gas phase, as extreme reactivity is expected to result from its strong curvature and flagrant violation of the "isolated pentagon rule"⁹. Even the less-curved fullerene C_{36} is so reactive that it spontaneously oligomerizes in the solid state^{10–12}.

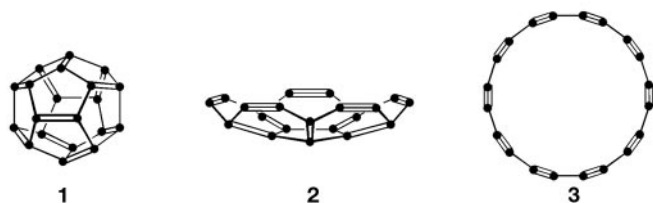


Figure 1 Isomers of C_{20} . 1, cage; 2, bowl; 3, ring.

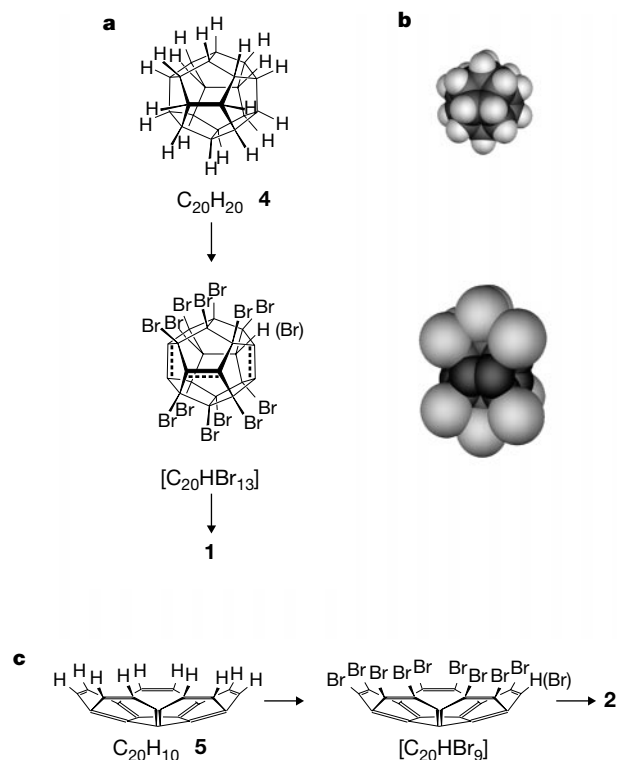


Figure 2 Preparation of precursors for cage 1 and bowl 2. **a**, Bromination of dodecahedrane 4. **b**, Space-filling models of $C_{20}H_{20}$ and $[C_{20}HBr_{13}]$ (see text). **c**, Bromination of corannulene 5.

An obvious starting point for the production of cage 1 is $C_{20}H_{20}$, dodecahedrane (4 in Fig. 2; refs 13, 14); quantities of this compound that are sufficient for synthetic purposes are now available through the highly optimized "isodrin-pagodane route"^{15,16}. To transform 4 into 1, however, the strongly bound hydrogen atoms must first be replaced by other, more weakly bound atoms, such as chlorine or bromine. Unfortunately, functionalization with sterically less demanding chlorine did not provide a good route to 1; cation mass-spectroscopic measurements on $C_{20}Cl_{16–20}$ samples did show elimination of chlorine atoms following electron impact ionization, but significant cage fragmentation was also observed¹⁴. Functionalization of 4 with more weakly bound bromine atoms seemed the way out of this dilemma. There was, however, a significant—and, for a long time, unsolved¹⁴—synthetic obstacle: the total (or at least near-total) substitution of the small hydrogen atoms by the voluminous bromine atoms failed due to the build-up of excessive molecular strain^{17,18}. The way to avoid this obstacle proved to be a 'brute-force' bromination protocol (see Methods), which reproducibly gave a product of average elemental composition $C_{20}HBr_{13}$ (denoted here as $[C_{20}HBr_{13}]$). This product was spectroscopically and chemically identified as primarily a multitude of isomeric $C_{20}H_{0–3}Br_{14–11}$ trienes, which—in spite of their strong deviation from planarity¹⁹ proved oxygen-insensitive (Fig. 2a). During the bromination process, elimination reactions provide relief from the steric overcrowding on the molecular periphery (Fig. 2b). Catalytic reduction of the bromination product back to 4 proved that the cage survived this treatment.

The cation mass spectrum of $[C_{20}HBr_{13}]$ thus prepared (Fig. 3) shows singly and doubly charged C_{20} ions with up to 14 bromine atoms attached. The relative intensities of the $C_{20}H_n^+$ ions (1.0: 0.4: 0.2: 0.1 for $n = 0, 1, 2, 3$) indicate only minor loss of the residual hydrogen atoms. The high intensity of the C_{20}^+ and C_{20}^{2+} ions and the very low abundance of smaller fragments attests to the remarkable kinetic stability of this C_{20} cluster.

In order to exclude the possibility that cage 1 might have isomerized into the (possibly more stable) bowl 2 during the course of the debromination, we pursued the independent generation of 2. This was likewise achieved by gas-phase debromination of a $[C_{20}HBr_9]$ precursor prepared by forcing bromination of corannulene 5 (ref. 20; see Fig. 2c and Methods). The 70-eV cation mass spectrum of this material closely resembles that of brominated dodecahedrane, showing C_{20}^+ clusters with up to nine bromine atoms attached (relative intensities of the $C_{20}H_n^+$ ions 1.0: 0.8: 0.8: 0.3 for $n = 0, 1, 2, 3$).

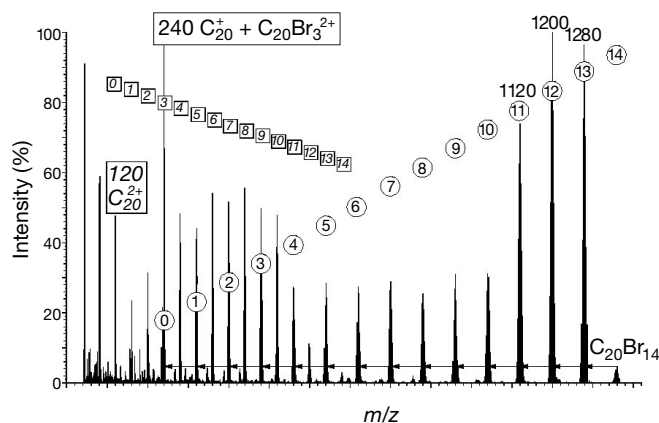


Figure 3 Cation mass spectrum of $[C_{20}HBr_{13}]$; 70-eV electrons were used for ionization. A series of singly charged clusters $C_{20}Br_n^+$ with $n = 0–14$ (circles), and a similar series of doubly charged clusters $C_{20}Br_n^{2+}$ (squares), can be distinguished. m/z , mass-to-charge ratio.

Electron-impact mass spectroscopy alone can neither distinguish between the C_{20} species derived from these two very different brominated hydrocarbons, nor even reveal whether or not they are still different isomers. To establish this, we converted the C_{20} species to their anions, and examined them by photoelectron spectroscopy (PES). In these experiments, a special ion source was used (see Methods) that produces an intense and very stable negative ion beam for several hours from a few milligrams of material. The anion mass spectra obtained for brominated dodecahedrane and brominated corannulene (Fig. 4) resemble those of the cation mass spectra, showing a sequence of brominated C_{20} clusters, with the number of attached bromine atoms varying from 13 to 0 in the first case and from 9 to 0 in the second. Again, the high kinetic stability of the C_{20} cluster is manifest.

Photoelectron spectra of mass-selected C_{20}^- clusters obtained from the two brominated precursors, and from graphite in a standard laser evaporation cluster source, are shown in Fig. 5. The spectrum in Fig. 5c is identical to published spectra that have been safely attributed to ring 3 on the basis of ion mobility measurements⁶. The spectra of Fig. 5a and b exhibit notably different features. Both have a reduced electron affinity with respect to ring 3, and show distinct vibrational patterns. Ring 3 has an electron affinity of 2.44 ± 0.03 eV, and exhibits a vibrational progression of $2,260 \pm 100$ cm^{-1} . The isomer generated from brominated corannulene has a significantly lower electron affinity of 2.17 ± 0.03 eV, but shows a similar vibrational progression of $2,060 \pm 50$ cm^{-1} . The isomer generated from brominated dodecahedrane has an electron affinity of 2.25 ± 0.03 eV, and exhibits a

vibrational progression of 730 ± 70 cm^{-1} ; at 0.27 eV above the first ionization threshold, another progression sets in, this one having a spacing of 260 ± 40 cm^{-1} . The spectra have been measured several times and showed no dependence on the source conditions. Thus, the PES data point to the presence of a single, unique C_{20} isomer in each case. Exact mass selection, however, is crucial; admixture of hydrogenated species $C_{20}H_1^-$ or $C_{20}H_2^-$ led to a blurring of the vibrational structure. The fact that all three anions can be generated selectively, moved over finite periods of time through various parts of the apparatus, separated from other ions, and then photoionized to the neutral carbon allotrope, demonstrates that all three C_{20} clusters, particularly the fullerene C_{20} , have a lifetime of at least the total flight time (0.4 ms).

The measured vibrational pattern of the bowl 2 is in qualitative agreement with predictions (G. Seifert, personal communication). As in ring 3, the additional electron occupies the lowest antibonding orbital of the triple bonds. Detachment of the electron will thus lead to a high-energy acetylenic stretching mode; the calculations predict frequencies $> 2,000$ cm^{-1} for such modes^{21,22}, which agree with the frequencies observed in both cases. For cage 1, the situation is much more complex; still, the observed 730 cm^{-1} progression is clearly more consistent with that expected for a closed 'polyolefinic' cage (1) than with the vibrational progression expected for any other open polycyclic isomer that has acetylenic edges. Unfortunately, it is

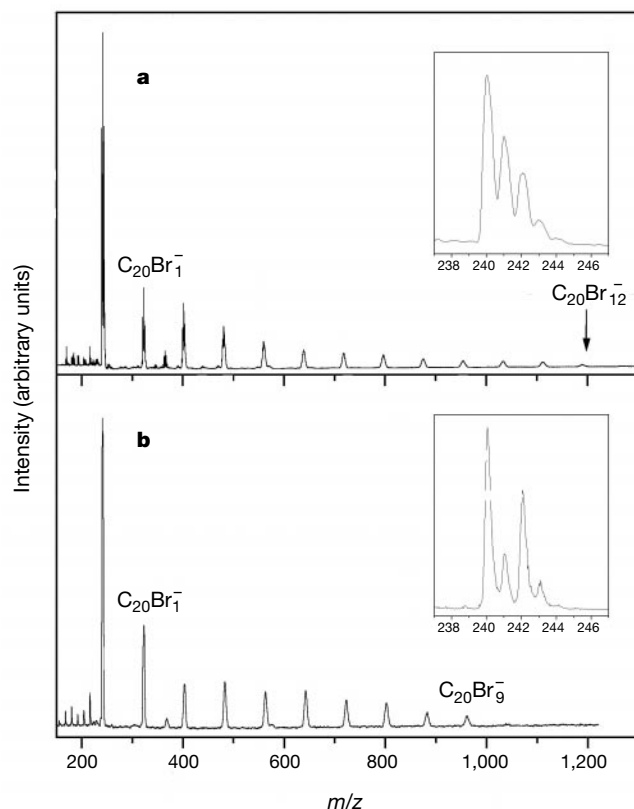


Figure 4 Anion mass spectra of $[C_{20}HBr_{13}]^-$ (a) and $[C_{20}HBr_9]^-$ (b). The molecules have been debrominated and charged by insertion into a gas discharge. In both cases we see a series of $C_{20}Br_n^-$ clusters, with n ranging from 13 to 0 in the upper spectrum (a) and from 9 to 0 in the lower spectrum (b). Each inset shows a magnification of the C_{20} mass peak. Note that due to the natural abundance of the ^{13}C isotope (1.1%), about 18% of the C_{20} intensity appears at mass 241.

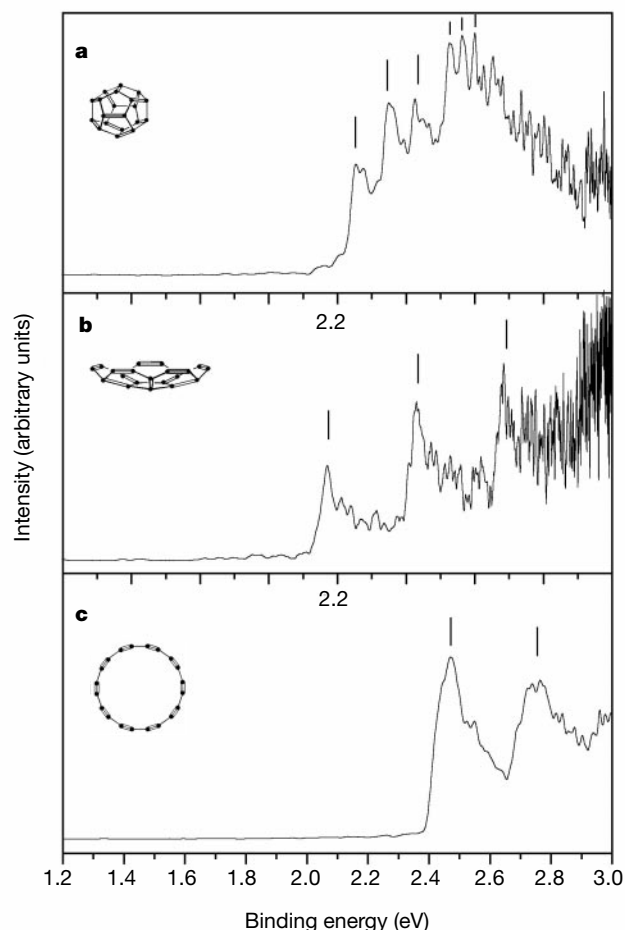


Figure 5 Photoelectron spectra of the mass-selected C_{20} clusters. a, C_{20}^- produced from $[C_{20}HBr_{13}]^-$. The wavelength of the detachment laser was 380 nm (3.26 eV). b, C_{20}^- produced from brominated corannulene, measured at 380 nm (3.26 eV). c, C_{20}^- clusters generated by a standard laser evaporation cluster source, measured at 308 nm (4.03 eV). A KrF excimer laser was used as the evaporation laser, and room-temperature helium as carrier gas.

not yet possible to make definite assignments, in part because the fourfold degeneracy of the highest occupied molecular orbital of the cage is avoided by a Jahn–Teller deformation away from perfect icosahedral symmetry, the nature of which is still disputed²². Knowledge of the ground-state symmetry of the neutral cage, however, is indispensable for the prediction of the vibrational modes excited upon electron detachment. We note that there is not even agreement on the relative heats of formation of **1** and **2** in spite of intensive numerical efforts over many years². The photoelectron spectrum of the unique C_{20} species derived from dodecahedrane thus stands as a benchmark test for quantum-mechanical methods. It is hoped that the experimental results reported here will stimulate further theoretical activities.

Until now fullerenes have been produced primarily by carbon condensation processes. The route to cage **1** that we report here is, to our knowledge, the first which makes use of a precursor with a rationally designed carbon core. □

Methods

Bromination procedures

Bromination of **4**: the solution of **4** in dry, deoxygenated bromine was refluxed under visible light irradiation for 3 d in a pressure flask, allowing the generated HBr to leak out. After extraction of the more hydrogen-rich components, the nearly insoluble, deep-red product was sublimed at 300 °C. Bromination of **5**: the solution of **5**, bromine and $FeCl_3$ in 1,1,2,2-tetrachloroethane was refluxed for 24 h.

Mass selection and PES characterization

The precursor substance was immobilized by a drop of toluene within a quartz tube, which was electrically heated to a temperature of 150–250 °C. A pulse of ultrapure helium flushed evaporated material into a copper tube, where a gas discharge was fired by applying a short high-voltage pulse to an isolated pin. The gas discharge fragmented the precursor molecules and negatively charged the products with a very high efficiency. After expansion into the vacuum, the ions entered a reflectron time-of-flight mass spectrometer²³, which could be used either to measure mass spectra of the ions or to insert ions of a given mass into a magnetic bottle photoelectron spectrometer. Here the ions were decelerated and irradiated by a pulsed laser beam. The velocities of the emitted electrons were measured and transformed into a kinetic-energy distribution. The energy resolution of the spectrometer is better than 20 meV for electron kinetic energies below 0.5 eV.

Received 4 October 1999; accepted 6 June 2000.

- Kroto, H. W. Smaller carbon species in the laboratory and in space. *Int. J. Mass Spectrom. Ion Process.* **138**, 1–15 (1994).
- van Orden, A. & Saykally, R. J. Small carbon clusters. Spectroscopy, structure, and energetics. *Chem. Rev.* **98**, 2313–2357 (1998).
- Yang, S. *et al.* UPS of 2–30-atom carbon clusters: Chains and rings. *Chem. Phys. Lett.* **144**, 431–436 (1988).
- Handschuh, H. *et al.* Stable configurations of carbon clusters: Chains, rings, and fullerenes. *Phys. Rev. Lett.* **74**, 1095–1098 (1995).
- Wakabayashi, T. *et al.* Photoelectron spectroscopy of C_n produced from laser ablated dehydroannulene derivatives having carbon ring size of $n = 12, 16, 18, 20$ and 24 . *J. Chem. Phys.* **107**, 4783–4787 (1997).
- Von Helden, G. *et al.* Do small fullerenes exist only on the computer? Experimental results on $C_{30}^{+/-}$ and $C_{24}^{+/-}$. *Chem. Phys. Lett.* **204**, 15–22 (1993).
- Heilbronner, E. & Dunitz, J. D. *Reflections on Symmetry* (VCH, Weinheim, 1993).
- Von Helden, G., Gotts, N. G. & Bowers, M. T. Experimental evidence for the formation of fullerenes by collisional heating of carbon rings in the gas phase. *Nature* **363**, 60–63 (1993).
- Kroto, H. W. The stability of the fullerenes C_n with $n = 24, 28, 32, 36, 50, 60$ and 70 . *Nature* **329**, 529–531 (1987).
- Piskoti, C., Yarger, J. & Zettl, A. C_{36} , a new carbon solid. *Nature* **393**, 771–774 (1998).
- Fowler, P. W. *et al.* C_{36} , a hexavalent building block for fullerene compounds and solids. *Chem. Phys. Lett.* **300**, 369–378 (1999).
- Koshio, A., Inakuma, M., Sugai, T. & Shinohara, H. A preparative scale synthesis of C_{36} by high-temperature laser vaporization: Purification and identification of $C_{36}H_6$ and $C_{36}H_6O$. *J. Am. Chem. Soc.* **122**, 398–399 (2000).
- McEwen, A. B. & Schleyer, P. v. R. In-plane aromaticity and trishomoaromaticity: A computational evaluation. *J. Org. Chem.* **51**, 4357–4368 (1986).
- Wahl, F., Würth, J. & Prinzbach, H. The pagodane route to dodecahedranes: An improved approach to the $C_{20}H_{20}$ parent framework: partial and total functionalizations—does C_{20} -fullerene exist? *Angew. Chem. Int. Edn Engl.* **32**, 1722–1726 (1993).
- Prinzbach, H. & Weber, K. From an insecticide to Plato's Universe—the pagodane route to dodecahedranes: New pathways and new perspectives. *Angew. Chem. Int. Edn Engl.* **33**, 2239–2257 (1994).
- Bertau, M. *et al.* From pagodanes to dodecahedranes—search for a serviceable access to the parent $C_{20}H_{20}$ hydrocarbon. *Tetrahedron* **53**, 10029–10040 (1997).
- Cioslowsky, J., Edgington, L. & Stefanov, B. B. Steric overcrowding in perhalogenated cyclohexanes, dodecahedranes, and [60]fullerenes. *J. Am. Chem. Soc.* **117**, 10381–10384 (1995).

- Beckhaus, H.-D. *et al.* Experimental enthalpies of formation and strain energies for the caged $C_{20}H_{20}$ pagodane and dodecahedrane frameworks. *J. Am. Chem. Soc.* **116**, 11775–11778 (1994); **117**, 8885 (1995).
- Melder, H.-P. *et al.* Unsaturated dodecahedranes - synthesis of the highly pyramidalized, highly reactive $C_{20}H_{18}$ and $C_{20}H_{16}$ olefins. *Res. Chem. Intermed.* **22**, 667–702 (1996).
- Scott, L. T. *et al.* Corannulene. A three-step synthesis. *J. Am. Chem. Soc.* **119**, 10963–10968 (1997).
- Galli, G., Gygi, F. & Golaz, J.-C. Vibrational and electronic properties of neutral and negatively charged C_{20} clusters. *Phys. Rev. B* **57**, 1860–1867 (1998).
- Dušekas, G. & Larsson, S. Bond lengths and reorganization energies in fullerenes and their ions. *Theor. Chem. Acc.* **97**, 110–118 (1997).
- Haberland, H., Kornmeier, H., Ludewigt, C., Risch, A. & Schmidt, M. A double/triple time-of-flight mass spectrometer for the study of photoprocesses in clusters, or how to produce cluster ions with different temperatures. *Rev. Sci. Instrum.* **62**, 2621–2625 (1991).

Acknowledgements

We thank J. Leonhardt, G. Leonhardt-Lutterbeck, S. Ruf and C. Warth for technical assistance, and G. Seifert, S. Larsson and R. C. Haddon for discussions. This work was supported by the BASF AG, the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and the US National Science Foundation.

Correspondence and requests for material should be addressed to H.P. (e-mail: horst.prinzbach@orgmail.chemie.uni-freiburg.de).

Chemical characterization of bohrium (element 107)

R. Eichler^{*†}, W. Brühl[‡], R. Dressler[§], Ch.E. Düllmann^{*†}, B. Eichler[†], H. W. Gäggeler^{*†}, K. E. Gregorich^{||}, D. C. Hoffman^{||}, S. Hübener[#], D. T. Jost[†], U. W. Kirbach^{||}, C. A. Laue^{||}, V. M. Lavanchy[†], H. Nitsche^{||}, J. B. Patin[†], D. Piguet[†], M. Schädel[‡], D. A. Shaughnessy^{||}, D. A. Strellis^{||}, S. Taut[#], L. Tobler[†], Y. S. Tsygancov^{**}, A. Türler^{*†}, A. Vahle[#], P. A. Wilk^{||} & A. B. Yakushev^{**}

^{*} *Departement für Chemie und Biochemie, Universität Bern, CH-3012 Bern, Switzerland*

[†] *Labor für Radio- und Umweltchemie, Paul Scherrer Institut, CH-5232 Villigen, Switzerland*

[‡] *Gesellschaft für Schwerionenforschung, D-64291 Darmstadt, Germany*

[§] *Institut für Analytische Chemie, Technische Universität Dresden, D-01062 Dresden, Germany*

^{||} *Nuclear Science Division, Lawrence Berkeley National Laboratory, California 94720, USA*

^{||} *Department of Chemistry, University of California at Berkeley, California 94720, USA*

[#] *Institut für Radiochemie, Forschungszentrum Rossendorf, D-01314 Dresden, Germany*

^{**} *Flerov Laboratory of Nuclear Reactions, Joint Institute for Nuclear Research, 141980 Dubna, Russia*

The arrangement of the chemical elements in the periodic table highlights resemblances in chemical properties, which reflect the elements' electronic structure. For the heaviest elements, however, deviations in the periodicity of chemical properties are expected^{1–3}: electrons in orbitals with a high probability density near the nucleus are accelerated by the large nuclear charges to relativistic velocities, which increase their binding energies and cause orbital contraction. This leads to more efficient screening of the nuclear charge and corresponding destabilization of the outer d and f orbitals: it is these changes that can give rise to unexpected chemical properties. The synthesis of increasingly heavy elements^{4–6}, now including that of elements 114, 116 and 118, allows the investigation of this effect, provided sufficiently long-lived isotopes for chemical characterization are available⁷. In the case of elements 104 and 105, for example, relativistic effects interrupt characteristic trends in the chemical properties of the elements constituting the corresponding columns of the periodic table⁸, whereas element 106 behaves in accordance with the

expected periodicity^{9–12}. Here we report the chemical separation and characterization of six atoms of element 107 (bohrium, Bh), in the form of its oxychloride. We find that this compound is less volatile than the oxychlorides of the lighter elements of group VII, thus confirming relativistic calculations¹³ that predict the behaviour of bohrium, like that of element 106, to coincide with that expected on the basis of its position in the periodic table.

The efforts of the Dubna and Darmstadt laboratories to synthesize element 107 in a nuclear fusion reaction of ⁵⁴Cr ions and a ²⁰⁹Bi target resulted in the first unambiguous identification of an isotope of element 107, ^{262m}Bh, in 1981 (ref. 14) where m indicates the meta-stable state. The half-life of 8 ms determined for the α -particle decay of the synthesized ^{262m}Bh nuclide, and even the half-life of 440 ms for the more stable isotope ²⁶⁴Bh (ref. 15), are too short for an investigation of chemical properties: the fastest radiochemical techniques allowing α -spectrometric identification of isolated nuclides currently require a half-life of about 5 s (ref. 7). Earlier searches for the presumably longer-lived, more neutron-rich Bh nuclides ²⁶⁶Bh and ²⁶⁷Bh, applying gas chemical separation techniques of volatile oxide and oxyhydroxide compounds of Bh, have failed^{16,17}. Recently, however, the nuclide ²⁶⁷Bh was discovered¹⁸ in the nuclear fusion reaction ²⁴⁹Bk(²²Ne, 4n). The production cross-section of ²⁶⁷Bh was 58^{+33}_{-15} pb at a beam energy of 117 ± 1 MeV. The measured half-life of 17^{+14}_{-6} s and its α -particle decay ($E_\alpha = 8.83 \pm 0.03$ MeV) make ²⁶⁷Bh an ideal nuclide for chemical investigations.

The systematic order in the periodic table suggests that Bh is a member of group VII and thus is chemically similar to its lighter homologues, technetium (Tc) and rhenium (Re). Investigations with short-lived nuclides ¹⁰⁸Tc ($T_{1/2} = 5.2$ s) and ¹⁶⁹Re ($T_{1/2} = 16$ s)^{19,20} showed that the most promising approach to test this assumption and characterize bohrium chemically was by studying the volatility of its oxychlorides using online gas chromatography²¹. According to these studies, works in preparative chemistry^{22–24} and mass spectrometric investigations^{25,26}, the formation of MO₃Cl (where M is Tc or Re) is most probable with oxidizing chlorinating gases, if only a few single metal atoms are available. With the generally high volatility of group VII oxychlorides, an excellent gas-phase separation from heavy actinides and lighter transactinides as well as from Pb, Bi and Po is possible. A good separation from the latter is very important as several nuclides of Pb, Bi and Po, which are formed as by-products in nuclear transfer reactions, severely interfere with the unambiguous detection of the nuclear decay of the investigated Bh isotope.

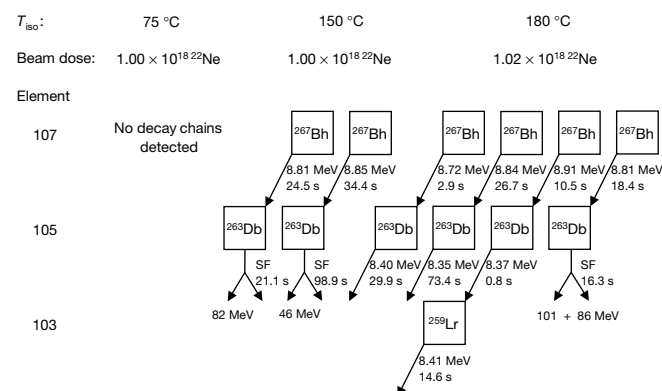


Figure 1 The six nuclear decay chains of ²⁶⁷Bh leading to ²⁶³Db and ²⁵⁹Lr. These were observed at isothermal temperatures of 180 °C and 150 °C, which allowed the unambiguous identification of bohrium after chemical separation as volatile BhO₃Cl. The α -particle and spontaneous fission (SF) decay energies are given in mega-electronvolts and the observed lifetimes in seconds. As no ²⁶⁷Bh was detected at an isothermal temperature of 75 °C and nearly equal beam doses of ²²Ne were applied to the target at all three temperatures, we were able to evaluate the volatility of BhO₃Cl quantitatively.

To assess the influence of relativistic effects on the chemical properties of BhO₃Cl, fully relativistic density-functional calculations have been performed for the group VII oxychlorides MO₃Cl, where M is Tc, Re or Bh¹³. The results of these calculations have shown that the electronic structure of BhO₃Cl is very similar to that of TcO₃Cl or ReO₃Cl. Increasing dipole moments and electric dipole polarizabilities in the group suggest a decreasing volatility in the sequence TcO₃Cl > ReO₃Cl > BhO₃Cl. In addition, classical extrapolations down the groups of the periodic table, using empirical correlations of thermochemical properties analogous to the predictions made for thermochemical properties of seaborgium compounds²⁷, predict BhO₃Cl to be more stable and less volatile than ReO₃Cl or TcO₃Cl (for full details, see http://wwwl.psi.ch/www_lch_hn/Bh_chemistry_prediction.pdf). Bh is thus expected to behave chemically like a typical member of group VII of the periodic table and the volatility of its compound BhO₃Cl is expected to be the lowest within the group.

A target of ²⁴⁹Bk (670 $\mu\text{g cm}^{-2}$) covered with a layer of ¹⁵⁹Tb (100 $\mu\text{g cm}^{-2}$) was prepared at LBNL on a thin (2.77 mg cm^{-2}) Be foil. At the Philips cyclotron of the Paul Scherrer Institute the target was irradiated for about four weeks with typically 1.6×10^{12} particles of ²²Ne per second at a beam energy in the middle of the target of 119 ± 1 MeV, producing ²⁶⁷Bh in the reaction ²⁴⁹Bk(²²Ne, 4n). ¹⁷⁶Re was simultaneously produced in the reaction ¹⁵⁹Tb(²²Ne, 5n) and served as a yield monitor for the chemical separation process. Nuclear reaction products, recoiling from the thin ²⁴⁹Bk target, were adsorbed onto the surface of carbon particles suspended in He gas and were then continuously transported through a thin capillary to the online gas chemistry apparatus at a distance of a few metres. There, the aerosol particles were collected on quartz wool inside the reaction oven kept at 1,000 °C. Reactive gases (HCl and O₂) were introduced in order to form volatile oxychlorides as well as to oxidize the carbon particles. The chromatographic separation took place downstream in the adjoining isothermal section of the quartz column (length 1.5 m, inner diameter 1.5 mm). The retention time of compounds in the column is mainly dependent on their adsorption interaction with the chlorinated column surface at a given isothermal temperature (T_{iso}) and on the carrier gas velocity. Only molecules formed with radionuclides having a longer nuclear lifetime than the chemical retention time pass the chromatography

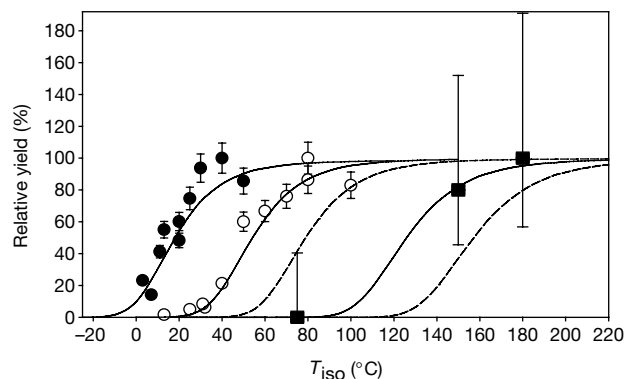


Figure 2 Results of isothermal gas adsorption chromatographic separations of group VII metal oxychlorides in quartz columns. The relative yields of the compounds ¹⁰⁸TcO₃Cl (filled black circles)²¹, ¹⁶⁹ReO₃Cl (open circles)²¹ and ²⁶⁷BhO₃Cl (filled black squares, our experimental results), after separation with the OLGA device, are shown as a function of isothermal temperature (T_{iso}). The error bars indicate a 68% confidence interval. The lines represent calculated relative yields applying a microscopic model of the adsorption process based on a Monte Carlo approach²⁸, with standard adsorption enthalpies of -51 kJ mol^{-1} (TcO₃Cl), -61 kJ mol^{-1} (ReO₃Cl) and -75 kJ mol^{-1} (BhO₃Cl), respectively. The dashed lines represent the calculated relative yield concerning the 68% confidence interval of the standard adsorption enthalpy of BhO₃Cl from -66 to -81 kJ mol^{-1} .

set-up and are subsequently attached to CsCl aerosol particles and transported to a detection system. The CsCl particles were deposited in vacuum on thin ($\sim 30 \mu\text{g cm}^{-2}$) polyethylene foils mounted on the circumference of a stepwise rotating wheel. Every 10 s the collected samples were successively moved between a series of 12 pairs of passivated ion-implanted planar silicon detectors to measure, in a time-resolved way, the energy of α -particles and spontaneous fission decays with about 70% detection efficiency. The nuclide ^{267}Bh decays by α -particle emission to ^{263}Db ($T_{1/2} = 27$ s), an isotope of element 105, which decays either by spontaneous fission (57%) or by α -particle emission to ^{259}Lr ($T_{1/2} = 6.1$ s), which either decays by a further α -particle emission (75%) or by spontaneous fission (25%). If we now observe a decay sequence whose α -particle energies and decay times are consistent with the expected decay chain, then the probability that this signal is indeed originating from a decay of ^{267}Bh is much higher than if we just observe single α -particle or spontaneous fission decays. The yield of ^{176}Re was measured with a high-purity germanium (HPGe) γ -ray detector. The overall yield of the complete separation process from the thermalization of the recoiling fusion products to the sample preparation in the detection device was on the average about 16% for ^{176}Re . Experiments were conducted at three different isothermal temperatures. Four correlated decay chains attributed to the decay of ^{267}Bh were detected at 180°C (see Fig. 1). Because at this isothermal temperature a small fraction of $^{212}\text{Pb}/^{212}\text{Bi}$ nuclides also passed the chromatography column, decays of their daughter nuclide ^{212}Po ($E_\alpha = 8.785$ MeV) partly obscured the detection of ^{267}Bh . Therefore, about 1.3 of the 4 observed decay chains were attributed to random correlations (N_R) unrelated to the decay of ^{267}Bh . At 150°C , 2 decay chains were observed ($N_R = 0.1$). Finally, at an isothermal temperature of 75°C , where $^{169}\text{ReO}_3\text{Cl}$ still passed through the isothermal part of the column with about 80% relative yield, no ^{267}Bh was registered, with a sensitivity similar to that at the two higher isothermal temperatures. Taking into account the contribution of random correlations, the relative yields of ^{267}Bh , observed at three different isothermal temperatures, were evaluated; we took the yield observed at an isothermal temperature of 180°C as 100% relative yield (see Fig. 2).

The adsorption properties of BhO_3Cl were quantified using a microscopic model of the adsorption process²⁸. This model yields the standard adsorption enthalpy (ΔH_{ads}) of the compound on the stationary phase, assuming a standard adsorption reaction on the quartz surface. The standard adsorption enthalpy of BhO_3Cl on the quartz surface was evaluated to be $-75^{+9}_{-6} \text{ kJ mol}^{-1}$ (68% c.i.), with $T_{1/2} = 17$ s for ^{267}Bh . This value compares well with the calculated value of $-78^{+5}_{-5} \text{ kJ mol}^{-1}$, obtained for BhO_3Cl applying a physisorption model adjusted to the experimental adsorption enthalpies of ReO_3Cl (ref. 13), as well as with $-74^{+12}_{-12} \text{ kJ mol}^{-1}$, deduced from thermodynamic extrapolations (see data at http://wwwl.psi.ch/www_lch_hn/Bh_chemistry_prediction.pdf). Hence, with a probability higher than 90%, BhO_3Cl exhibits a stronger adsorption interaction with the chlorinated quartz surface than TcO_3Cl ($-51^{+3}_{-3} \text{ kJ mol}^{-1}$) and ReO_3Cl ($-61^{+3}_{-3} \text{ kJ mol}^{-1}$)²⁰. We obtained additional evidence for stronger similarity of BhO_3Cl to ReO_3Cl than to TcO_3Cl experimentally. In model experiments TcO_3Cl was too volatile to be adsorbed on CsCl aerosols after chemical separation. TcO_3Cl could only be transported to the counting system with FeCl_2 aerosol particles, presumably owing to their reducing surface²⁰. BhO_3Ce was transported with CsCl aerosols, strongly supporting the results of our isothermal chromatography experiments.

An empirical correlation between the standard adsorption enthalpy on quartz and the standard sublimation enthalpy, established for chlorides and oxychlorides¹¹, allows us to estimate the standard sublimation enthalpy of BhO_3Cl as $89^{+21}_{-18} \text{ kJ mol}^{-1}$, compared to the values of $49^{+12}_{-12} \text{ kJ mol}^{-1}$ and $66^{+12}_{-12} \text{ kJ mol}^{-1}$ deduced in

the same manner for TcO_3Cl and ReO_3Cl , respectively. With the unambiguous identification of Bh after chemical separation, it is clear that Bh, in the form of a volatile oxychloride compound (presumably BhO_3Cl), behaves like a typical member of group VII of the periodic table. As expected from relativistic calculations of molecular properties and from classical extrapolations of periodic trends within the groups of the periodic table, Bh shows a lower volatility of its oxychloride compound than its Re or Tc homologues. □

Received 15 May; accepted 31 July 2000.

- Pyykkö, P. Relativistic effects in structural chemistry. *Chem. Rev.* **88**, 563–594 (1988).
- Pershina, V. G. Electronic structure and properties of the transactinides and their compounds. *Chem. Rev.* **96**, 1977–2010 (1996).
- Schwerdtfeger, P. & Seth, M. in *Encyclopaedia Of Computational Chemistry* Vol. 4, 2480–2499 (Wiley, New York, 1998).
- Oganessian, Yu., Ts. et al. Synthesis of nuclei of the superheavy element 114 in reactions induced by ^{48}Ca . *Nature* **400**, 242–245 (1999).
- Oganessian, Yu., Ts. et al. Synthesis of superheavy nuclei in the $^{48}\text{Ca} + ^{244}\text{Pu}$ reaction. *Phys. Rev. Lett.* **83**, 3154–3157 (1999).
- Ninov, V. et al. Observation of superheavy nuclei produced in the reaction of ^{86}Kr with ^{208}Pb . *Phys. Rev. Lett.* **83**, 1104–1107 (1999).
- Gäggeler, H. W. in *Proceedings of The Robert A. Welch Foundation, 41st Conference on Chemical Research, The Transactinide Elements* 43–63 (The Robert A. Welch Foundation, Houston, 1997).
- Kratz, J. V. in *Heavy Elements and New Phenomena* Ch. 4 (eds Greiner, W. & Gupta, R. K.) 129–193 (World Scientific, Singapore, 1999).
- Schädel, M. et al. Chemical properties of element 106 (seaborgium). *Nature* **388**, 55–57 (1997).
- Zvara, I. et al. Chemical identification of element 106 (thermochemistry of oxochlorides). *Radiochim. Acta* **81**, 179–187 (1998).
- Türler, A. et al. First measurement of a thermochemical property of a seaborgium compound. *Angew. Chem. Intl Edn Engl.* **38**, 2212–2213 (1999).
- Lougheed, R. Oddly ordinary seaborgium. *Nature* **388**, 21–22 (1997).
- Pershina, V. & Bastug, T. The electronic structure and properties of group 7 oxychlorides, MO_3Cl , where M = Tc, Re, and element 107, Bh. *J. Chem. Phys.* **113**, 1441–1446 (2000).
- Münzenberg, G. et al. Identification of element 107 by α correlation chains. *Z. Phys. A* **300**, 107–108 (1981).
- Hofmann, S. et al. The new element 111. *Z. Phys. A* **350**, 281–282 (1995).
- Zvara, I. et al. Experimental approach to the chemical identification of element 107 as eka-rhenium. II. Search for spontaneously fission isotopes of element 107 in the products of the reaction $^{249}\text{Bk} + ^{22}\text{Ne}$ by the method of chromatography. *Sov. Radiochem.* **26**, 72–76 (1984); translated from *Radiochim. Acta* **26**, 76–82 (1984).
- Schädel, M. et al. Radiochemical search for neutron-rich isotopes of nielsbohrium in the $^{16}\text{O} + ^{254}\text{Es}$ reaction. *Radiochim. Acta* **68**, 7–12 (1995).
- Wilk, P. A. et al. Evidence for New Isotopes of Element 107: ^{266}Bh and ^{267}Bh . *Phys. Rev. Lett.* (in the press).
- Eichler, R. et al. The gas phase oxide and oxyhydroxide chemistry of trace amounts of rhenium. *Radiochim. Acta* **87**, 151–159 (1999).
- Eichler, R. et al. Gas phase chemistry of technetium and rhenium oxychlorides. *Radiochim. Acta* **88**, 87–93 (2000).
- Gäggeler, H. W. et al. OLGA II, an on-line gas chemistry apparatus for applications in heavy element research. *Nucl. Instrum. Meth. A* **309**, 201–208 (1991).
- Feller, J., Oppermann, H., Binnewies, M. & Milke, E. Zum chemischen Transport von Rhenium und Rheniumoxiden. *Z. Naturforsch.* **53b**, 184–190 (1998).
- Drobot, D. V., Korshunov, B. G. & Aleksandrova, V. A. Partial triangulation of the phase diagram for the rhenium-chlorine-oxygen system. *Russian J. Inorg. Chem.* **17**, 914–916 (1972); translated from *Zh. Neorg. Khim.* **17**, 1762–1766 (1972).
- Guest, A. & Lock, C. J. L. Technetium chlorides and oxochlorides. *Can. J. Chem.* **50**, 1807–1810 (1972).
- Gibson, J. K. Synthesis and mass spectrometry of halide and oxide halide (F and Cl) vapour species of technetium, tantalum and tungsten. *J. Fluor. Chem.* **55**, 299–311 (1991).
- Rinke, K., Klein, M. & Schäfer, H. Massenspektrometrische Untersuchungen an Halogeniden des Rheniums und Technetiums. *J. Less-Comm. Met.* **12**, 497–503 (1967).
- Eichler, B., Türler, A. & Gäggeler, H. W. Thermochemical characterization of seaborgium compounds in gas adsorption chromatography. *J. Phys. Chem. A* **103**, 9296–9306 (1999).
- Zvara, I. Simulation of thermochromatographic processes by the Monte Carlo method. *Radiochim. Acta* **38**, 95–101 (1985).

Acknowledgements

We are indebted to the Office of Science, Office of Basic Energy Research, Division of Chemical Sciences, of the US Department of Energy, for making the ^{249}Bk target material available through the transplutonium element production program at the Oak Ridge National Laboratory. We thank the staff of the PSI Philips cyclotron for providing intense beams of ^{22}Ne . This work was supported in part by the Office of Science, Office of High Energy and Nuclear Physics, Division of Nuclear Physics, of the US Department of Energy and by the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to H.W.G. (e-mail: gaeggeler@iac.unibe.ch). Further information may be obtained from http://wwwl.psi.ch/www_lch_hn/Bh_chemistry_prediction.pdf.

Mid-depth recirculation observed in the interior Labrador and Irminger seas by direct velocity measurements

Kara L. Lavender*, Russ E. Davis* & W. Brechner Owens†

* Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0230, USA

† Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

The Labrador Sea is one of the sites where convection exports surface water to the deep ocean in winter as part of the thermohaline circulation. Labrador Sea water is characteristically cold and fresh, and it can be traced at intermediate depths (500–2,000 m) across the North Atlantic Ocean, to the south and to the east of the Labrador Sea^{1–3}. Widespread observations of the ocean currents that lead to this distribution of Labrador Sea water have, however, been difficult and therefore scarce. We have used more than 200 subsurface floats to measure directly basin-wide horizontal velocities at various depths in the Labrador and Irminger seas. We observe unanticipated recirculations of the mid-depth (~700 m) cyclonic boundary currents in both basins, leading to an anticyclonic flow in the interior of the Labrador basin. About 40% of the floats from the region of deep convection left the basin within one year and were rapidly transported in the anticyclonic flow to the Irminger basin, and also eastwards into the subpolar gyre. Surprisingly, the float tracks did not clearly depict the deep western boundary current, which is the expected main pathway of Labrador Sea water in the thermohaline circulation. Rather, the flow along the boundary near Flemish Cap is dominated by eddies that transport water offshore. Our detailed observations of the velocity structure with a high data coverage suggest that we may

have to revise our picture of the formation and spreading of Labrador Sea water, and future studies with similar instrumentation will allow new insights on the intermediate depth ocean circulation.

In the North Atlantic, surface waters of the meridional overturning circulation transport heat to high latitudes, where deep convection forms intermediate and deep waters in localized regions.

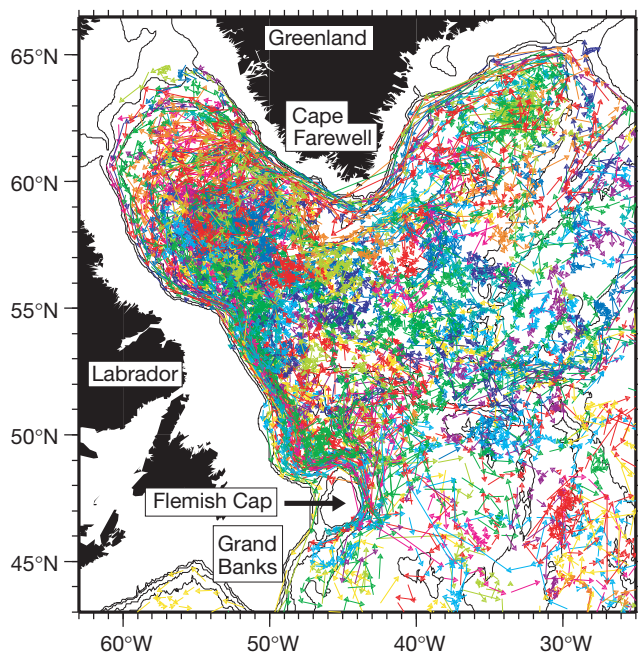


Figure 1 Trajectories of floats in the Labrador and Irminger seas at nominal depths of 400, 700 and 1,500 m. Each float is represented by one colour, and each arrow represents one subsurface drift cycle. Gaps between arrows represent the time the float was at the surface. Data shown are from November 1994 to the end of April 1999, representing over 200 years of data. Contours mark the 500-m, 1,000-m, 2,000-m, 3,000-m and 4,000-m isobaths.

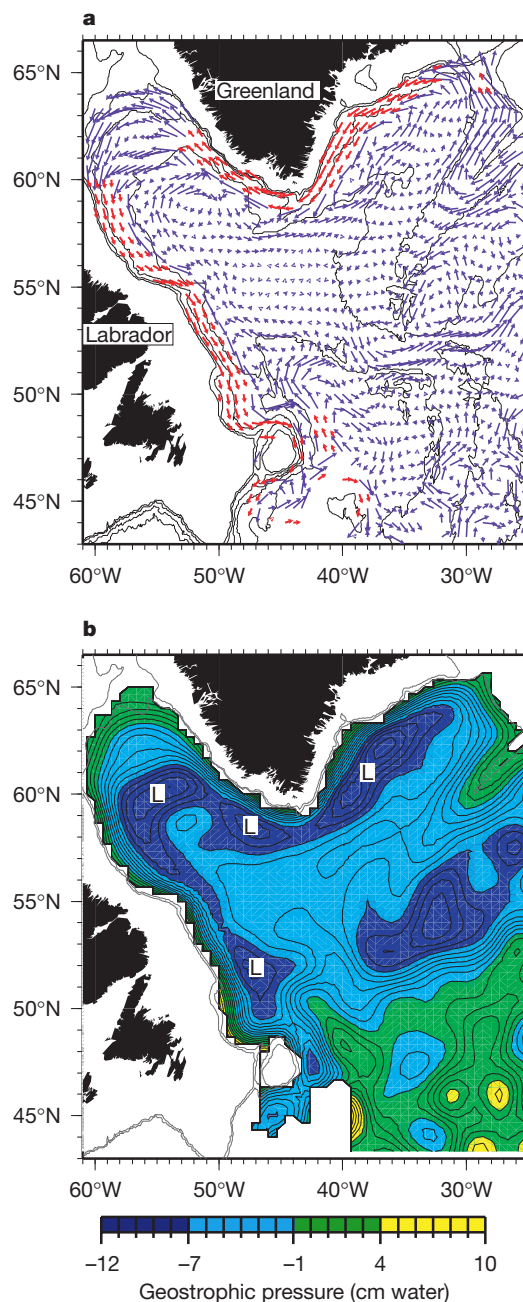


Figure 2 Objectively mapped mean circulation at 700 m depth. **a**, Velocity shown as displacement vectors. Blue arrows indicate distance travelled over 30 days for speeds $< 5 \text{ cm s}^{-1}$; red arrows indicate distance travelled over 8 days for speeds $\geq 5 \text{ cm s}^{-1}$. Contours mark isobaths as in Fig. 1. **b**, Geostrophic pressure measured in centimetres of water (black contours). 'L's mark selected low-pressure centres, and grey contours mark the 500-m and 1,000-m isobaths. To construct these maps, velocity data from all drift depths were corrected to 700 m using geostrophic shears from the Levitus climatology^{20,21}, and were averaged in bins roughly 100 km square. These data were then objectively mapped using a gaussian covariance function with a decorrelation length scale of 185 km and noise derived from the variance of the velocity field.

Long-term variability in the formation process^{4,5}, and subsequent spreading of these water masses, may alter the meridional heat flux, a central component of the climate system⁶. Labrador Sea water (LSW), located between depths of 500 and 2,000 m, is formed in winter in the interior of the basin where cold, dry, offshore winds force convective overturning. Although hydrographic data have identified locations of deep convection⁷ and have inferred a cyclonic general circulation in the Labrador Sea⁸, direct velocity measurements have been limited to point measurements from a few moored current meters^{9,10}. Consequently the residence time and pathways of LSW within the basin, and its exit routes into the North Atlantic, have mostly been inferred from indirect measurements.

As part of the World Ocean Circulation Experiment (WOCE) and the Labrador Sea Deep Convection Experiment¹¹, more than 200 neutrally buoyant subsurface P-ALACE floats¹² and SOLO floats¹³ were deployed in the northern North Atlantic. The floats drift at nominal depths of 400 m, 700 m, or 1,500 m for preset periods between 3.5 and 20 days. At the end of its cycle time, each float ascends to the sea surface to transmit data via the Argos satellite system. The floats report position data, as well as vertical profiles of ocean temperature and salinity measured upon ascent or descent.

Since November 1994, more than 7,400 profiles of temperature and salinity and more than 200 years (cumulative) of drift-velocity data have been collected (Fig. 1). The drift-velocity data were used to objectively map the mean circulation at 700 m depth (Fig. 2). The strongest flows in the mean field are along the coasts of Greenland (East and West Greenland currents) and Labrador (Labrador Current), with speeds reaching 12 cm s^{-1} . The North Atlantic Current meanders eastwards at $3\text{--}7 \text{ cm s}^{-1}$ from northeast of Flemish Cap, the topographic feature centred at 47.5° N , 45° W . The most notable feature is the unexpected anticyclonic counter-current to the interior of the boundary currents. This anticyclonic flow is comprised of a series of recirculations that are visible as closed regions of cyclonic flow in Fig. 2a, and as low-pressure centres marked with an 'L' in Fig. 2b. This flow has not been described in circulation regimes inferred from hydrography^{8,9},

although northward flow to the interior of the Labrador Current was measured with current meters at one location⁷.

The circulation patterns described above appear at all three drift depths and at all times of year. The flow in the upper 1,000 m is generally faster than that below, particularly in the boundary currents and in the anticyclonic flow northeast of Flemish Cap, where speeds weaken at depth by up to 5.5 cm s^{-1} . Although seasonal variability in the boundary currents could not be resolved, the recirculation to the north of Flemish Cap is faster in winter (January–April) than in summer (July–October), whereas the reverse is true of the recirculation in the western Irminger basin. The Flemish Cap recirculation appears to be fed by the Labrador Current, whose seasonal variability has been linked to variations in wind forcing¹⁰.

Mixed-layer depths computed from temperature profiles for the winters of 1997 and 1998 (Fig. 3) show that convectively mixed layers deeper than 400 m were scattered across the basin interior, but the deepest mixed layers ($> 800 \text{ m}$) were concentrated in the western basin. This augments the picture found from a hydrographic survey¹¹ in February–March 1997. The floats also measured recently formed deep mixed layers as late as mid-April, and in regions outside the western basin. Mixed layers in 1998 were shallower than in 1997, indicative of the interannual variability in LSW formation⁵.

Changes in heat content between consecutive vertical profiles of temperature provide a measure of the air–sea heat flux. The greatest computed average winter heat flux ($-345 \pm 95 \text{ W m}^{-2}$) occurred in the northern third of the basin, whereas the central portion containing the deepest mixed layers had an average winter heat flux of only $-49 \pm 46 \text{ W m}^{-2}$. Therefore the circulation within the western basin must be an important factor in determining where deep convection occurs.

Previous work⁷ suggests that a winter cyclonic gyre in the western Labrador basin traps surface waters in a region that is repeatedly exposed to cold offshore winds. The resulting heat loss to the atmosphere then sustains deep convection. To investigate the existence of a localized gyre, the mean velocity field was mapped

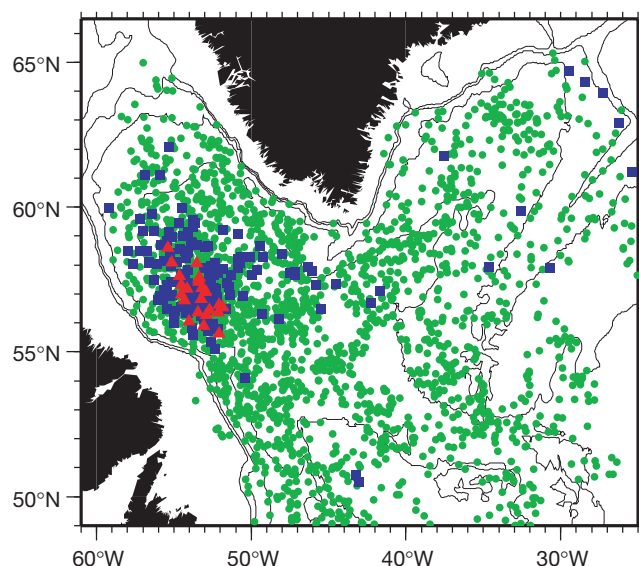


Figure 3 Winter mixed-layer depth for profiles measured in winters 1996–97 and 1997–98. The mixed-layer depth (MLD) is defined by the break between a least-squares fit to the upper layer, and a second-order polynomial plus exponential fit to the lower layer, of each temperature profile. $\text{MLD} \leq 400 \text{ m}$ are marked with green circles, $400 < \text{MLD} \leq 800 \text{ m}$ with blue squares, and $\text{MLD} > 800 \text{ m}$ with red triangles. Contours mark isobaths as in Fig. 1. The deepest mixed layers are located in the western basin, although mixed layers deeper than 400 m are found outside this region as well.

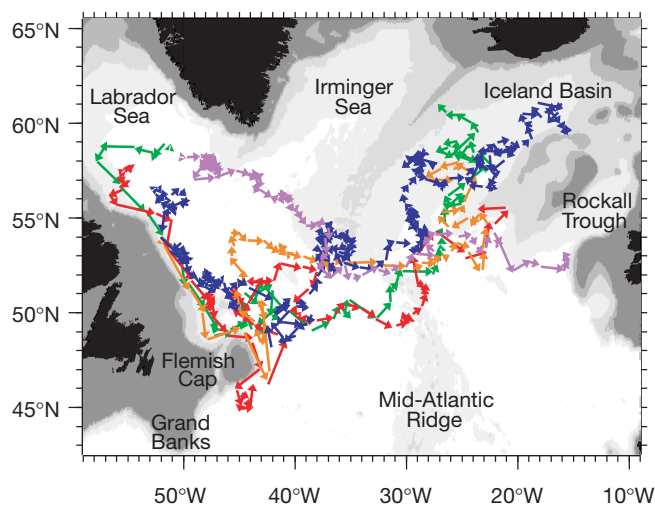


Figure 4 Trajectories of five floats that travelled from the central Labrador Sea to the eastern North Atlantic. Red, orange and green arrows indicate three floats drifting at 700 m depth; blue and purple arrows indicate two floats drifting at a depth of 1,500 m. From light to dark grey, shading indicates bathymetry $< 3,000 \text{ m}$, $< 2,000 \text{ m}$, $< 1,000 \text{ m}$ and $< 500 \text{ m}$. One float (purple) left the Labrador Sea in the unexpected anticyclonic current south of Greenland and arrived at Rockall Trough after only 2.8 years. The other four floats drifted in an eastward flow fed by the Labrador Current, with travel times comparable to previous estimates of spreading rates.

from winter data only (not shown). Nearly all of the deepest mixed layers were located within a weak (1.8 cm s^{-1}) cyclonic recirculation, centred near 56.5°N between the Labrador Current and the interior flow. This gyre is, however, masked by variability, suggesting that the important factor is the generally weak flow in the region. Floats that recorded the deepest convection remained in the region an average of 34.5 days, and in some cases up to 110 days, before measuring the deep mixed layers. There may be a localized winter cyclonic gyre in the convection region, but (more importantly) the sluggish flow retains water there for an extended period of time, during which the water experiences a large net heat loss that drives progressively deeper convection.

From the formation region, two main routes exist for LSW to leave the Labrador Sea. The first is an eastward outflow between 51°N and 54°N , fed by the Labrador Current. Three floats drifting at 700 m and one float drifting at 1,500 m followed this path to the eastern North Atlantic (Fig. 4). The floats' travel times compare well with spreading estimates inferred from tracers³. Floats at 700 m depth reached the Iceland basin in 1.6–2.6 years, at an average subsurface speed (including small-scale features) of 8.7 cm s^{-1} , while the 1,500-m-deep float arrived in 3.8 years, at an average subsurface speed of 6.6 cm s^{-1} .

The second pathway out of the Labrador Sea is in the anticyclonic countercurrent flowing from the interior of the basin into the Irminger Sea. Evidence that LSW travels directly to the Irminger basin is given by high concentrations of chlorofluorocarbons (CFCs) at intermediate depths^{3,14}, indicative of recently ventilated water. One 1,500-m-deep float (purple in Fig. 4) left the Labrador Sea in this countercurrent, and reached the eastern edge of the North Atlantic basin at Rockall trough in 2.8 years (average speed of 4.9 cm s^{-1}). The travel time via this alternative route across the basin is 1–1.5 years faster than previous estimates from tracers³.

LSW properties have also been traced south along the western boundary into the subtropical gyre^{1,15}, presumably having travelled in the deep western boundary current (DWBC). Surprisingly, none of the 1,500-m floats sampled the DWBC south of Flemish Cap (Fig. 5). The flow near Flemish Cap is strongly baroclinic, with the southward DWBC located inshore of the 4,000-m isobath¹⁵ and below the northward, surface-intensified North Atlantic Current (NAC). At the surface, some floats were carried offshore by the

NAC, while others drifted onshore towards the DWBC region. At depth, a number of floats sampled an expected southward flow inshore of the 3,000-m isobath north of $\sim 47^\circ \text{N}$, with average measured speeds of up to 24.7 cm s^{-1} . Near Flemish Cap, however, a number of floats underwent cross-isobath flow with comparable speeds. The flow at the northern edge of Flemish Cap turns back to the northwest, feeding a recirculation of the interior anticyclonic countercurrent, while eastward and southeastward flow is observed throughout much of the region. The existence of these cross-isobath flows at the slope of Flemish Cap indicates a rapid exchange between the boundary current and offshore waters, and is largely responsible for the lack of floats travelling southward into the subtropical gyre. Measurements¹⁶ of high eddy kinetic energy east of Flemish Cap suggest that these flows are eddy-driven.

Floats that travelled along the western boundary and recirculated to the northwest near Flemish Cap did not measure significant variations in temperature and salinity at depths of 1,000–1,500 m, indicating this is a pathway of LSW. Most floats travelling east of $\sim 42^\circ \text{W}$, however, encountered the warmer and saltier water of the NAC at these depths. Hydrographic data¹⁷ show a deepening of isopycnals to the east of Flemish Cap, suggesting that LSW is still present in this region but below the depth of the floats (1,500 m). Thus both deep pathways described above transport LSW away from the western boundary.

Although floats did not drift south along the western boundary beyond $\sim 45^\circ \text{N}$, tracer studies^{1,17} clearly indicate that LSW does reach the subtropical gyre, presumably in the DWBC. One reason that the floats did not measure the boundary current may be high variability in the flow field near Flemish Cap, which might make intermittent the southward flow of LSW in the DWBC¹⁷. Another possibility is that LSW reaches the subtropical gyre further east, in recirculations on either side of the Mid-Atlantic Ridge^{18,19}. The questions of how and when intermediate and deep waters enter the subtropical gyre, and whether or not this exchange could be shut off by climate variability, are essential to understanding the long-term behaviour of the thermohaline circulation. These questions cannot be answered with the present observations. □

Received 2 March; accepted 22 June 2000.

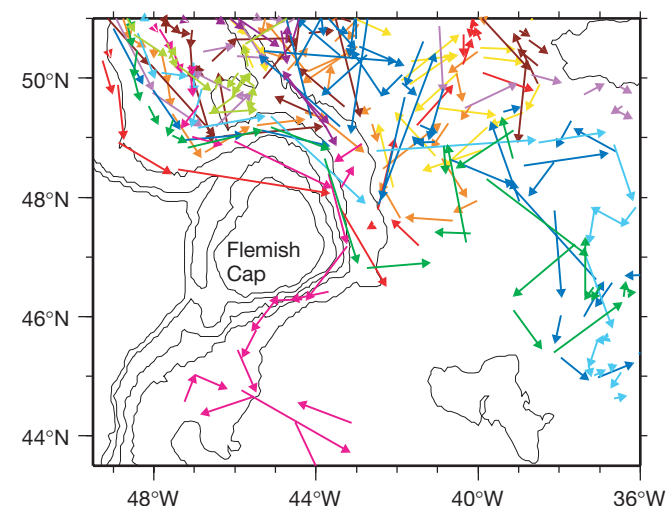


Figure 5 Displacements of floats drifting at 1,500 m depth near Flemish Cap. Colour scheme and bathymetric contours as in Fig. 1. There is no indication of a coherent southward flow along the boundary south of Flemish Cap. Instead deep, cross-isobath flows are observed, which provide a rapid exchange between the boundary and offshore waters.

1. Talley, L. D. & McCartney, M. S. Distribution and circulation of Labrador Sea Water. *J. Phys. Oceanogr.* **12**, 1189–1205 (1982).
2. Cunningham, S. A. & Haine, T. W. N. Labrador Sea Water in the Eastern North Atlantic. Part I: A synoptic circulation inferred from a minimum in potential vorticity. *J. Phys. Oceanogr.* **25**, 649–665 (1995).
3. Sy, A. *et al.* Surprisingly rapid spreading of newly formed intermediate waters across the North Atlantic Ocean. *Nature* **386**, 675–679 (1997).
4. Dickson, R. From the Labrador Sea to global change. *Nature* **386**, 649–650 (1997).
5. Dickson, R., Lazier, J., Meincke, J., Rhines, P. & Swift, J. Long-term coordinated changes in the convective activity of the North Atlantic. *Prog. Oceanogr.* **38**, 241–295 (1996).
6. Bryden, H. L. in *Interactions Between Global Climate Subsystems, The Legacy of Hann* (eds McBean, G. A. & Hantel, M.) 65–75 (American Geophysical Union, Washington DC, 1993).
7. Clarke, R. A. & Gascard, J.-C. The formation of Labrador Sea Water. I. Large-scale processes. *J. Phys. Oceanogr.* **13**, 1764–1778 (1983).
8. Lazier, J. R. N. The renewal of Labrador Sea Water. *Deep-Sea Res.* **20**, 341–353 (1973).
9. Clarke, R. A. Transport through the Cape Farewell–Flemish Cap section. *Rapp. P.-v. Réun. Cons. Int. Explor. Mer.* **185**, 120–130 (1984).
10. Lazier, J. R. N. & Wright, D. G. Annual velocity variations in the Labrador Current. *J. Phys. Oceanogr.* **23**, 659–678 (1993).
11. Lab Sea Group. The Labrador Sea Deep Convection Experiment. *Bull. Am. Meteorol. Soc.* **79**, 2033–2058 (1998).
12. Davis, R. E., Webb, D. C., Regier, L. A. & Dufour, J. The autonomous Lagrangian circulation explorer (ALACE). *J. Atmos. Oceanic Technol.* **9**, 264–285 (1992).
13. Davis, R. E., Sherman, J. T. & Dufour, J. Profiling ALACEs and other advances in autonomous subsurface floats. *J. Atmos. Oceanic Technol.* (submitted).
14. Min, D. H. *Studies of Large-scale Intermediate and Deep Water Circulation and Ventilation in the North Atlantic, South Indian and Northeast Pacific Oceans, and in the East Sea (Sea of Japan), using Chlorofluorocarbons as Tracers*. Thesis, Univ. California, San Diego (1999).
15. Pickart, R. S., Smethie, W. M., Lazier, J. R. N., Jones, E. P. & Jenkins, W. J. Eddies of newly formed upper Labrador Sea Water. *J. Geophys. Res.* **101**, 20711–20726 (1996).
16. Carr, M.-E. & Rossby, T. Pathways of transport in the North Atlantic Current from surface drifters and subsurface floats. *J. Geophys. Res.* (submitted).
17. Pickart, R. S., Spall, M. A. & Lazier, J. R. N. Mid-depth ventilation in the western boundary current system of the sub-polar gyre. *Deep-Sea Res.* **44**, 1025–1054 (1997).
18. Lozier, M. S. Evidence for large-scale eddy-driven gyres in the North Atlantic. *Science* **277**, 361–364 (1997).

19. Paillet, J., Arhan, M. & McCartney, M. S. Spreading of Labrador Sea Water in the eastern North Atlantic. *J. Geophys. Res.* **103**, 10223–10239 (1998).
20. Levitus, S., Burgett, R. & Boyer, T. P. *World Ocean Atlas 1994* Vol. 3, *Salinity* (NOAA Atlas NESDIS 3, US Department of Commerce, Washington, DC, 1994).
21. Levitus, S. & Boyer, T. P. *World Ocean Atlas 1994* Vol. 4, *Temperature* (NOAA Atlas NESDIS 4, US Department of Commerce, Washington DC, 1994).

Acknowledgements

We thank J. Dufour, J. Sherman, J. Valdes, R. Tavares and B. Guest for technical development and preparation of the floats, D. Newton and C. Wooding for help with data processing, and the numerous scientists who deployed the floats, often in adverse conditions at sea. This work was supported by the National Science Foundation, the Office of Naval Research, and the National Oceanic and Atmospheric Administration.

Correspondence and requests for materials should be addressed to K.L.L.
(e-mail: klavender@ucsd.edu)

Earthquakes as beacons of stress change

Leonardo Seeber & John G. Armbruster

Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA

Aftershocks occurring on faults in the far-field of a large earthquake rupture can generally be accounted for by changes in static stress on these faults caused by the rupture^{1,2}. This implies that faults interact, and that the timing of an earthquake can be affected by previous nearby ruptures^{3–6}. Here we explore the potential of small earthquakes to act as ‘beacons’ for the mechanical state of the crust. We investigate the static-stress changes resulting from the 1992 Landers earthquake in southern California which occurred in an area of high seismic activity stemming from many faults. We first gauge the response of the regional seismicity to the Landers event with a new technique, and then apply the same method to the inverse problem of determining the slip distribution on the main rupture from the seismicity. Assuming justifiable parameters, we derive credible matches to slip profiles obtained directly from the Landers mainshock^{7,8}. Our results provide a way to monitor mechanical conditions in the upper crust, and to investigate processes leading to fault failure.

The mechanical state of a fault can be characterized by Coulomb stress⁹ $CS = \tau - \mu(\sigma_n - p)$, a scalar, where τ is shear stress, σ_n is normal stress, μ is the coefficient of friction, and p is the pore pressure. In general, CS is not known, but a ‘beacon’ fault will experience a static stress change $\Delta CS = \Delta\tau - \mu'\Delta\sigma_n$ from a known shear dislocation on a causative fault¹⁰ (the ‘rupture’). In a linear isotropic elastic medium ΔCS is a scalar. In this simple formulation, the effective friction parameter μ' includes a possible change in pore pressure, but is independent of space and time. $\Delta CS \ll CS$ for beacons sufficiently far from the rupture (7.5 km in this case; Fig. 1) and only the component of $\Delta\tau$ parallel to the beacon’s slip vector is consequential. $\Delta CS > 0$ is expected to ‘encourage’ the beacon fault toward failure; $\Delta CS < 0$ is expected to ‘discourage’ it from failing. This expectation can be tested for earthquakes that are resolved as shear failures with specific location, strike, dip and rake. In general, ΔCS is significant within a rupture-length of any rupture¹¹. Both deformation and stress changes, however, are dominated by the largest ruptures¹². We consider the classical interaction experiment where ΔCS stems from a large rupture and affects a multitude of beacon earthquakes that are assumed to be too small to mutually interact¹.

The 1992 Landers sequence in southern California provides the causative rupture in this experiment: it is dominated by the 28 June,

$M_W = 7.2$ mainshock and includes the 23 April, $M_L = 6.1$ Joshua Tree ‘foreshock’ and the 28 June, $M_L = 6.5$ Big Bear ‘aftershock’ (Fig. 1). The rupture is within a broad zone of seismicity associated with the plate boundary. A set of relocated and quality-selected focal mechanisms for most $M \geq 1.5$ earthquakes in this zone since 1980 have been interpreted for fault structure and preferred nodal planes (data are available at <ftp://sccc.gps.caltech.edu/pub/focal/focall.nano>). This seismicity stems from many relatively small, scattered faults ranging widely in orientation and kinematics which serve as ideal beacons for stress change¹³ (Fig. 1). ΔCS is calculated for each beacon as the compound effect of 28 elemental dislocations¹⁰ approximating the rupture⁷.

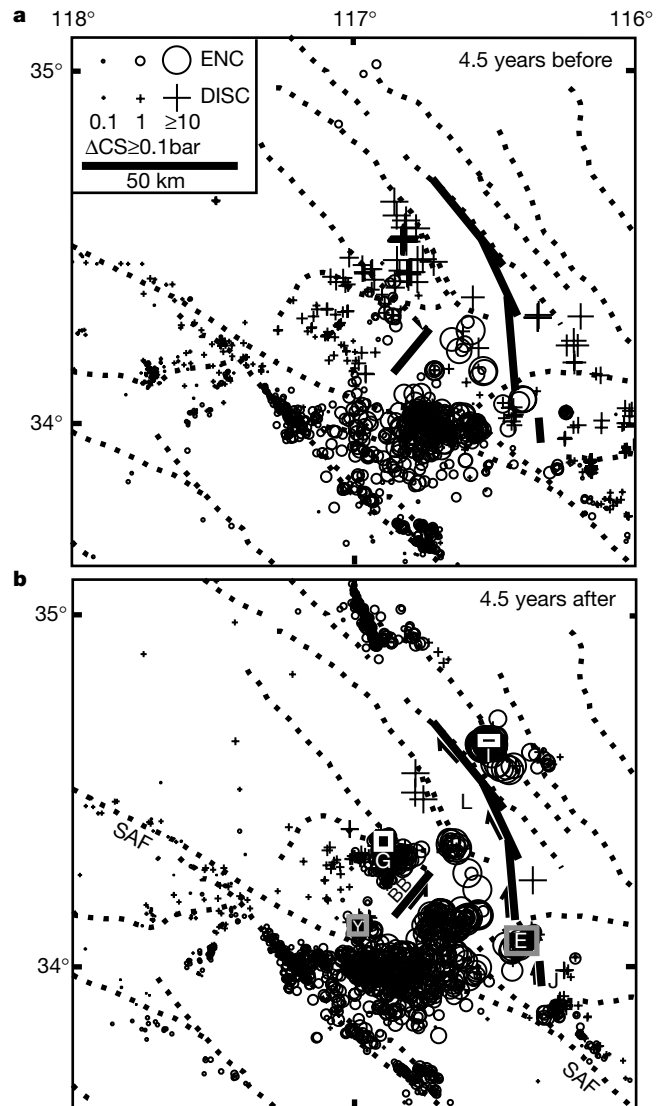


Figure 1 Earthquakes as beacons of the static stress change from the 23 April–28 June 1992 composite rupture. Heavy lines^{7,9}: J, Joshua; L, Landers; BB, Big Bear. Coulomb stress change calculated for each beacon is either encouraging (ENC, $\Delta CS \geq 0.1$ bar; circles) or discouraging (DISC, $\Delta CS < 0.1$ bar; crosses). Beacons after the rupture (1,680 encouraged and 538 discouraged earthquakes) are expected to occur if encouraged, while beacons before the rupture (681 encouraged and 553 discouraged earthquakes) provide the reference (for example, discouraged seismicity is turned off by the rupture). Excluded are beacons less than 7.5 km from the rupture, where non-elastic interactions may occur. Friction $\mu' = 0.8$. Early discouraged earthquakes are concentrated in four clusters (boxed) where local stress interactions are suspected: G, Gold Mountain; Y, Yucaipa; E, Eureka Park; I, Iron Ridge. SAF, San Andreas fault.

Earthquakes before the rupture (Fig. 1a) are not affected by it, but their ΔCS provide a reference: the numbers of encouraged (E) and discouraged (D) pre-rupture earthquakes are similar ($E/D \approx 1$). After the rupture, E/D is 5–10 times higher, primarily because of an increase in E (Fig. 2). A suppression of the discouraged pre-rupture earthquakes by the rupture is apparent in the overall response of the far field (Fig. 2c) and by comparing the spatial distribution of the discouraged and encouraged pre-rupture earthquakes in the near field before and after the rupture^{14,15} (Fig. 1). Although a broad range of friction values ($0.4 < \mu' < 1.0$) yields an E/D response qualitatively consistent with the Coulomb failure model, we obtain the strongest E/D response for a high value of friction⁹, $\mu' \approx 0.8$. This optimal friction value seems robust because it remains within ± 0.1 of 0.8 for several non-overlapping subsets of beacons during the first 1.5 years after the rupture. The apparent agreement between these and other experimental values is surprising because μ' includes any Δp effect¹⁶.

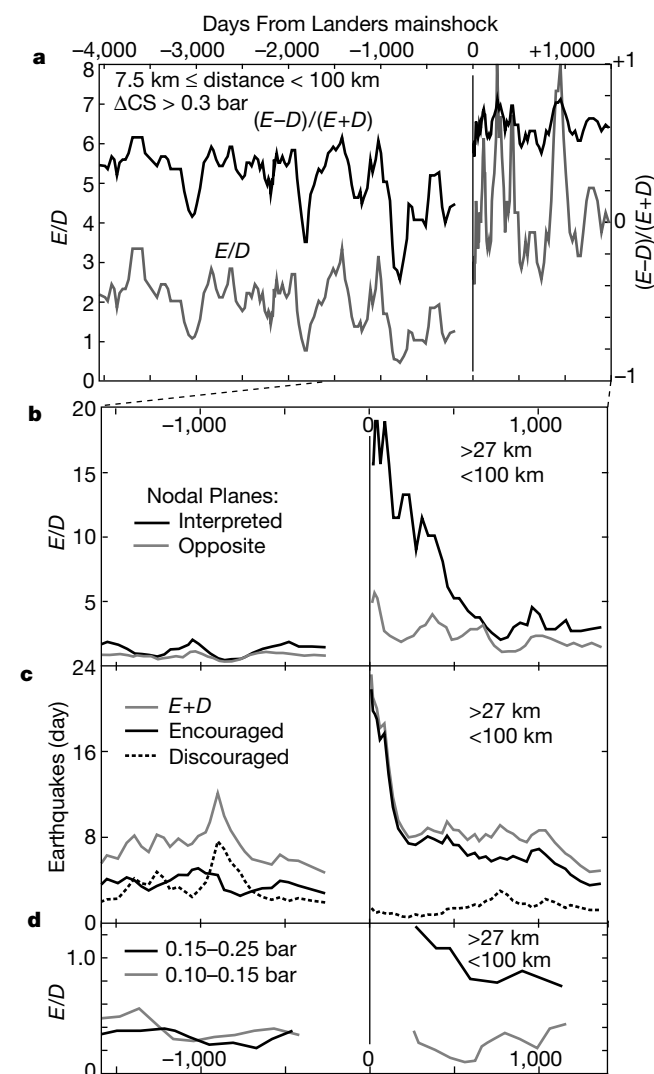


Figure 2 Histograms of ΔCS effects on seismicity from the Landers rupture⁷. The effect is gauged by the numbers of encouraged (E) versus discouraged (D) beacon earthquakes. Friction $\mu' = 0.8$ yields the optimal E/D response and is used in all calculations. **a**, E/D and $(E-D)/(E+D)$ as alternative response parameters for all beacons more than 7.5 km from the rupture. **b**, E/D response of distant beacons (more than 27 km from the ruptures; $\Delta CS \geq 0.2$ bar) according to the nodal planes interpreted to represent the faults, or the opposite planes. **c**, The same beacons as in **b** separated into E , D and $E+D$. **d**, Two subsets of the beacons in **b** within narrow adjacent ranges of ΔCS just above and below the 0.15-bar response threshold for these data.

Earthquakes are very sensitive stress beacons; a significant increase of E/D is detected for $0.15 \leq \Delta CS < 0.25$ bar (ref. 2; Fig. 2d). They also retain a memory of the rupture long afterwards. In the far-field, E/D is largest within a few months of the rupture; it decreases gradually afterwards, but it is still about twice the average pre-rupture level several years after the rupture (Fig. 2b). In contrast, the overall rate of seismicity in the far field (for example, $E+D$ in Fig. 2c) returns to pre-rupture levels within a year. The far-field effect of the rupture persists as a change in the kinematics of seismogenic faulting, but not necessarily as a rate increase.

According to the Coulomb failure model, some encouraged faults rise above critical stress and are expected to fail immediately, whereas discouraged faults are pushed away from failure and will take longer to reach it. The post-rupture deficit in discouraged earthquakes (Fig. 2c) is expected to persist until tectonics has re-loaded all negative ΔCS faults to their pre-rupture stress level. Tectonic loading is slow, particularly for secondary faults. Even for the San Andreas fault, the encouraging ΔCS effect southwest of the Landers rupture is equivalent to a decade of tectonic loading^{4,5}. Thus, a simple Coulomb formulation can account for a prolonged E/D response. The high rate of encouraged earthquakes, however, is also persistent after the mainshock (Fig. 2c). Thus failure is delayed for encouraged as well as discouraged earthquakes, in violation of a simple Coulomb model⁹. A delay in the nucleation of earthquakes can be accounted for by fluid flow in response to Δp (poroelastic response)¹⁵ and/or by rate-and-state-dependent friction¹⁷.

In spite of data and model limitations, Fig. 2 and many similar tests demonstrate some success in modelling the beacon response to a known large rupture. We expand on these experiments by turning the problem around and extracting information on the rupture from the seismicity. We parametrize the approximately 75 km long Landers rupture (Fig. 1; overlapping segments are combined) as 26 segments, each 3 km long, 15 km deep, and with uniform slip. Total

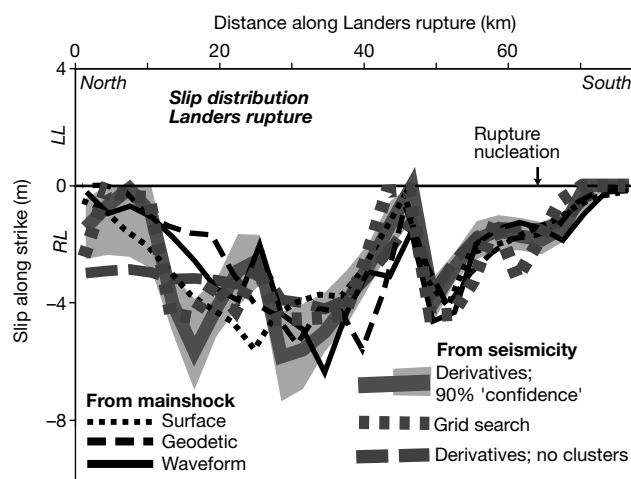


Figure 3 Comparison of slip distributions for the 1992 Landers rupture. Data derived from subsequent seismicity (thick and grey) and directly from the mainshock (thin and black; dotted, measured at the surface⁸; dashed, modelled from geodetic data⁸; solid, modelled from waveform and geodetic data⁸; RL, right-lateral; LL, left-lateral). Random sampling of beacons by a bootstrap technique produces a '90% confidence' range. The range of solutions is broader when different parameters and iteration techniques are tested. In these examples, slip is either traded between one segment and all others and is optimized by a grid search (dotted), or is determined according to the derivatives of the score with respect to slip (solid and dashed; only RL slip allowed; weighting is an increasingly weak function of ΔCS through the iteration). Beacons from four clusters outlined in Fig. 1 are excluded in the dashed profile. All iterations start from a uniform slip distribution.

moment is fixed and starting slip values are the same for all segments. The Big Bear and Joshua ruptures are assigned fixed uniform slip⁹. We then iterate for the slip distribution that yields the optimal response, as predicted by elastic interaction and Coulomb failure. E/D tends to be unstable where data are scarce (Fig. 2a). This optimal value is therefore chosen as the maximum value of $(E - D)$ integrated over a post-rupture period (3 years in Fig. 3). Beacons are weighted positively with ΔCS . Each step in an iteration maximizes the score by redistributing slip on the main rupture while keeping total moment constant and computes a new ΔCS for each beacon. This process is repeated until the score cannot be further improved. We have tested many iterations covering a range of starting, weighting and declustering parameters, spatial and temporal constraints on the data, and iteration schemes. We could clearly differentiate iteration schemes that yielded inconsistent spiky results from others that yielded relatively consistent and credible results.

In one class of iterations, slip change ΔS on one segment is traded with opposite change $-\Delta S/25$ on each of the other 25 segments, in each step. $\Delta S = K$ (0, or 1, or -1) is determined by a grid search for the highest score. After all segments have had this role, the procedure is repeated until the score has reached the maximum for a particular value of K . Then the process can be repeated for a smaller K , and so on. We generally start these iterations with $\Delta S = 2$ m and stop them at $\Delta S = 0.02$ m. Even when K is held constant, iterations of this kind tend to converge rapidly on permissible slip distributions with persistent features (see 'grid search', dotted line in Fig. 3). In another class of iterations, one segment trades slip with only one of the others in each step, and then with each of the others in turn. As in the previous case, this procedure is repeated for all segments, and then again in subsequent sweeps until improvements in the score for each sweep become vanishingly small. These iterations are particularly laborious because they sample the solution space in great detail over slip distributions with extreme variations; they may yield high scores, but they tend to converge on unrealistically spiky and inconsistent slip distributions. The need for imposing a degree of smoothing is a common feature of inversion problems¹⁸; in this case it may derive in part from clustering in the field of beacons. In a third class of iterations, all segments are treated equally in each step. The change in slip for a segment ΔS_i is proportional to how much it affects the score: $\Delta S_i = C[d(\text{score})/dS_i]$. The factor C is chosen from a set of values by grid-search for the optimal score. The final task in the n th iteration step is to normalize the absolute values of slip on each segment to maintain the total moment invariant: $S_{in} = [S_{i(n-1)} + \Delta S_{in}][\sum_j |S_{j(n-1)}|/\sum_j |S_{j(n-1)} + \Delta S_{jn}|]$. Iterations of this kind (see 'derivatives', solid and dashed lines in Fig. 3) tend to converge more rapidly than iterations by 'grid search'. In spite of fundamental differences, these methods tend to yield similar slip distributions (for example, Fig. 3).

The weighting function $w(\Delta CS)$ is one of the most important factors in the outcome of the iterations. In most iterations the weight was antisymmetric about $\Delta CS = 0$. Generally, the weight drops gradually with $|\Delta CS|$ (in Fig. 3: dotted line, by 0.01 per bar; solid and dashed lines, decreasing to 0.01 per bar) until $|\Delta CS| \approx 0.15$ bar (in Fig. 3: dotted line, $w = 0.1$ at $|\Delta CS| = 0.1$ bar; solid and dashed lines, $w = 0.2$ at $|\Delta CS| = 0.2$ bar) and then drops rapidly to zero for smaller $|\Delta CS|$ (in Fig. 3: dotted line, 1 per bar; solid and dashed lines, increasing to 2 per bar). This kink in the weight function correlates with the threshold value of ΔCS (Fig. 2d). The weight may change during the iteration (only dashed and solid in Fig. 3). At the beginning of the iteration the kink at the stress threshold is subtle because the slip distribution is uniform and the ΔCS values are far from realistic. At this stage, all beacons need to play a role in properly directing the convergence. As the iteration converges to more realistic ΔCS values, the weighting function is progressively more kinked. The shape of the weight function at low ΔCS affects the distance range over which beacons are effective; it is

particularly important for the northern part of the slip distribution which is controlled by a few clusters (Fig. 1).

Near-field beacons (7.5–27 km) are critical for resolving the main features of the slip distribution (10–20 km wavelengths), but they have a relatively weak E/D response early after the rupture because of dense clusters of discouraged earthquakes (Figs 1 and 2a versus Fig. 2b). The data were declustered by eliminating beacons sufficiently close in time, space and fault kinematics (and reduced by 1/5 in Figs 1, 2 and 3). Nevertheless, local interactions may still dominate the response for earthquakes in the Yucaipa, Gold Mountain, Eureka Peak and Iron Ridge clusters (Y, G, E and I in Fig. 1) which coincide with foci of early post-rupture moment release neglected in our rupture model ($M_L = 4.7$ at Y; $M_L = 5.3$ at G; $M_L = 5.6$ at E and 23 cm of left-lateral slip at I). Local interactions may be more prevalent where ΔCS is large and beacons are dense. In general, changes in declustering parameters have small effects on slip distribution. But when clusters of beacons suspected of violating the assumption of mutual independence are excluded, slip distributions tend to be less peaked (Fig. 3, dashed versus solid), probably because these clusters are relatively close to the rupture (Fig. 1). Temporal data selection is less critical. Progressively later batches of seismicity, as late as a year after the rupture, tend to yield comparable slip distributions.

Postseismic deformation related to the rupture, such as post-rupture creep⁸ and viscous response below the rupture²⁰, may significantly affect the beacon response. Tectonic loading is probably a relatively minor factor within a few years of the mainshock. Focal mechanisms are selected according to quality^{14,15} and are expected to introduce noise, but little systematic bias. Our structural interpretation of focal mechanisms for fault planes^{14,15} is another obvious source of noise, but it is substantially better than random, as demonstrated by the much subdued beacon effect from the opposite planes (Fig. 2b). Future experiments will benefit from continuing improvements in data quality and quantity. Beacons are more abundant around the southern part of the rupture where, not surprisingly, slip distribution features appear to be more robust (Fig. 3). Although $\mu' = 0.8$ yields optimal E/D , a uniform friction value is unrealistic in view of possible dependency on fault slip rate¹⁷, on presence of fluid, and on depth (temperature and pressure)¹⁶. Some of the early post-rupture earthquakes may have been discouraged and reflect inadequacy of the Coulomb formulation. Earthquakes sufficiently close to failure may occur in spite of a negative ΔCS , according to state-rate friction^{17,19} and an initial discouraging stress change may become encouraging as a result of fluid flow in poroelastic conditions¹⁶. Finally, alternative models of the rupture derived from the mainshock are significantly different⁸ (Fig. 3), so some of the post-rupture discouraged earthquakes may reflect inaccuracies with the rupture model (for example, Fig. 2).

We have obtained slip distributions of the Landers 1992 rupture from off-rupture aftershocks assuming a simple Coulomb interaction model and using different methods and assumptions. They compare with each other and with distributions obtained directly from the mainshock nearly as well as these direct determinations compare with each other (Fig. 3). Although small earthquakes contribute little to total deformation, their distribution in space, time, and fault kinematics is very sensitive to stress changes and offer unique opportunities to monitor stress. When applied to abundant high-resolution earthquake data, the approach outlined here is expected to provide new insight into a variety of phenomena affecting the mechanical state of the crust. Examples might include natural phenomena, such as fault creep or magma injection, or large engineering operations, such as impounding of water reservoirs or flooding of oil fields. Delay in the response of seismicity to known stress changes can be revealing about the nucleation of small earthquake ruptures. Conversely, mechanical changes resolved from seismicity may reveal processes preparatory to large earthquakes. □

Received 19 November 1999; accepted 19 July 2000.

- Stein, R. S. & Lisowski, M. The 1979 Homestead Valley earthquake sequence, California: Control of aftershock and postseismic deformation. *J. Geophys. Res.* **88**, 6477–6490 (1983).
- Reasenber, P. A. & Simpson, R. W. Response of regional seismicity to the static stress change produced by the Loma Prieta earthquake. *Science* **255**, 1687–1690 (1992).
- Hudnut, K. W., Seeber, L. & Pacheco, J. Cross-fault triggering in the November 1987 Superstition Hills earthquake sequence, southern California. *Geophys. Res. Lett.* **16**, 199–202 (1989).
- Jaumé, S. C. & Sykes, L. R. Changes in the state of stress on the southern San Andreas Fault resulting from the California earthquake sequence of April to June 1992. *Science* **258**, 1325–1328 (1992).
- Harris, R. A. & Simpson, R. W. Changes in static stress on southern California faults after the 1992 Landers earthquake. *Nature* **360**, 251–254 (1992).
- Stein, R. S. The role of stress transfer in earthquake occurrence. *Nature* **402**, 605–609 (1999).
- Wald, D. J. & Heaton, T. H. Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bull. Seismol. Soc. Am.* **84**, 668–691 (1994).
- Hudnut, K. W. *et al.* Seismic displacement of the 1992 Landers Earthquake sequence. *Bull. Seismol. Soc. Am.* **84**, 625–645 (1994).
- King, G. C. P., Stein, R. S. & Lin, J. Static stress change and the triggering of earthquakes. *Bull. Seismol. Soc. Am.* **84**, 935–953 (1994).
- Erickson, L. L. *A Three-dimensional Dislocation Program with Applications to Faulting in the Earth*. Thesis, Stanford Univ. (1986).
- Hardebeck, J. L., Nazareth, J. J. & Hauksson, E. The static stress change triggering model: Constraints from two southern California aftershock sequences. *J. Geophys. Res.* **103**, 24427–24437 (1998).
- Deng, J. & Sykes, R. L. Evolution of the stress field in southern California and triggering of moderate-size earthquakes: A 200-year perspective. *J. Geophys. Res.* **102**, 9859–9886 (1997).
- Seeber, L. & Armbruster, J. G. The San Andreas fault system through the Transverse Ranges as illuminated by earthquakes. *J. Geophys. Res.* **100**, 8285–8310 (1995).
- Toda, S., Stein, R. S., Reasenber, P. A., Dieterich, J. H. & Yoshida, A. Stress transferred by the 1995 $M_w = 6.9$ Kobe, Japan shock: Effect on aftershocks and future earthquake probabilities. *J. Geophys. Res.* **103**, 24543–24565 (1998).
- Gross, S. & Burgmann, R. Rate and state of background stress estimated from the aftershocks of the Loma Prieta, California earthquake. *J. Geophys. Res.* **103**, 4915–4927 (1998).
- Scholz, C. *The Mechanics of Earthquakes and Faulting* 53–66, 328–330 (Cambridge Univ. Press, New York, 1990).
- Dieterich, J. H. & Kilgore, B. Implications of fault constitutive properties for earthquake prediction. *Proc. Natl Acad. Sci.* **93**, 3787–3794 (1996).
- Menke, W. *Geophysical Data Analysis: Discrete Inverse Theory* 53 (Academic Press, 1989).
- Harris, R. A. & Simpson, R. W. Suppression of large earthquakes by stress shadows: A comparison of Coulomb and state-and-rate failure. *J. Geophys. Res.* **103**, 24439–24451 (1998).
- Deng, J., Hudnut, K., Gurnis, M. & Hauksson, E. Stress loading from viscous flow in the lower crust and triggering of aftershocks following the 1994 Northridge, California, earthquake. *Geophys. Res. Lett.* **26**, 3209–3212 (1999).

Acknowledgements

We are grateful for helpful suggestions and/or discussions from J. Hardebeck, P. Reasenber, J. Vermilye, C. Scholz, D. Simpson, L. Sykes, R. Stein, B. Shaw, B. Menke and K. Jacob. Financial support was provided by the Southern California Earthquake Center, the US Geological Survey and the NSF.

Correspondence and requests for materials should be addressed to L.S.
(e-mail: nano@ldeo.columbia.edu).

The earliest known sauropod dinosaur

Eric Buffetaut*, Varavudh Suteethorn†, Gilles Cuny‡, Haiyan Tong*, Jean Le Loeuff§, Sasidhorn Khansubha† & Sutee Jongautcharyakul†

* Centre National de la Recherche Scientifique, 16 cour du Liégat, 75013 Paris, France

† Geological Survey Division, Department of Mineral Resources, Rama VI Road, Bangkok 10400, Thailand

‡ Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK

§ Musée des Dinosaures, 11260 Espérance, France

Sauropods were a very successful group of dinosaurs during the Jurassic and Cretaceous periods, but their earlier history is poorly known. Until now, the earliest reported sauropod bones were from the Early Jurassic^{1–3}, and the only tentative evidence of earlier sauropods was in the form of controversial footprints^{4,5}. Here we report the discovery of an incomplete sauropod skeleton from the Late Triassic period of Thailand, which provides the first osteological evidence of pre-Jurassic sauropods. This dinosaur is markedly different from prosauropods and substantiates

theoretical predictions that there was a fairly long period of sauropod evolution during the Triassic.

Dinosauria Owen, 1842

Saurischia Seeley, 1888

Sauropodomorpha Huene, 1932

Sauropoda Marsh, 1878

Isanosaurus attavipachi gen. et sp. nov.

Etymology. Generic name from *Isan*, the local name for north-eastern Thailand, and *sauros*, Greek for lizard. Specific name in honour of P. Attavipach, former Director General of the Thai Department of Mineral Resources, a long-time supporter of palaeontological research.

Holotype. Associated skeletal elements (Fig. 1) consisting of one cervical, one dorsal and six caudal vertebral centra, the neural arch of a posterior dorsal vertebra, two chevron bones, fragmentary ribs, a right sternal plate, a right scapula and a left femur (Palaeontological collection, Department of Mineral Resources, Thailand: CH4).

Horizon and locality. The bones were found in 1998 in a natural outcrop of dark red sandstones of the Nam Phong Formation at Phu Nok Khian hill near Ban Non Thaworn village, in Chaiyaphum Province, on the Khorat Plateau of northeastern Thailand. They are clearly remnants of a single skeleton that had largely been eroded away before the first elements were discovered. Unfused neurocentral sutures indicate that the individual (possibly about 6.5 m long) may not have been fully grown.

Age. The fluvial Nam Phong Formation contains palynomorphs showing that it cannot be younger than Rhaetian⁶, and overlies the Huai Hin Lat Formation, which is well dated as Norian on the basis of its vertebrate fauna and palynoflora. It is therefore well dated as late Norian or Rhaetian⁶. The only vertebrate fossil hitherto reported from the Nam Phong Formation was fused ischia referred to a large prosauropod⁷. Whether they might in fact belong to *I. attavipachi* cannot be ascertained because no ischia were found at Phu Nok Khian.

Diagnosis. A primitive sauropod dinosaur with a robust femur bearing a very prominent, acuminate, S-shaped fourth trochanter located in the proximal half of the bone.

Some characters of *I. attavipachi* clearly place it among the Sauropoda, whereas others indicate a basal position within that group. It has been compared with prosauropods⁸, especially the somewhat sauropod-like Melanosauridae^{9,10}, and with primitive sauropods. Although other primitive sauropods are known¹¹, comparisons were made mainly with the following sufficiently well described forms, with significant skeletal elements also present in the Thai specimen: *Vulcanodon karibaensis*¹ (basal Jurassic, Zimbabwe), *Barapasaurus tagorei*² and *Kotasaurus yamanpalliensis*³ (from the supposedly Early Jurassic Kota Formation of India, which may in fact be as recent as Early Cretaceous on the basis of palynology; G. V. R. Prasad, personal communication), *Zizhongosaurus chuanchengensis*¹² and *Gongxianosaurus shibeensis*¹³ (Early Jurassic, China), and *Shunosaurus lii*¹⁴ (Middle Jurassic, China). The vertebrae differ from those of prosauropods, but in many respects are less advanced than those of later sauropods. A short cervical centrum, with parapophyses at mid-height (suggesting a posterior position), is markedly opisthocoeleous, unlike the amphicoeleous centra of prosauropods or the anteriorly flat ones of *Gongxianosaurus*¹³. It shows a strong ventral median ridge, a primitive feature in sauropods¹⁵. Its sides are deeply concave rather than excavated by real pleurocoels as in more advanced sauropods. Such lateral depressions also occur on a posterior dorsal centrum. In this regard, the presacral vertebrae of *I. attavipachi* resemble those of *B. tagorei* and *S. lii*. An isolated neural arch, probably from a posterior dorsal vertebra, is remarkably tall, as in some later sauropods, but unlike the relatively low neural spines of prosauropods. However, the spine is longer (rostrocaudally) than it is broad (transversely), which is primitive for sauropods¹⁶. In this respect,

Isanosaurus is less advanced than *B. tagorei* and *Z. chuanchengensis*, in which the spine is broadened transversely; it resembles Middle Jurassic sauropods, such as *S. lii*, *Lapparentosaurus madagascariensis* from Madagascar and *Volkheimeria chubutensis* from Patagonia, which show laterally flattened dorsal neural spines^{15,17}. Incipient posterolateral laminae and a ridge extending from the transverse process to the base of the spine are sauropod features¹⁵ not seen in prosauropods (including sauropod-like forms such as *Lessemsaurus*¹⁷, from the Late Triassic of Argentina), and more developed in later forms¹⁵, including *Volkheimeria* and *Lapparentosaurus*¹⁷. The caudal centra are amphicoelous. The incomplete scapula is reminiscent of *Shunosaurus*, with a moderate, dorsally rounded proximal expansion (primitive for sauropods¹⁵), and a slight distal expansion. The semicircular sternal plate bears a low ridge on the outer surface.

The 76-cm-long femur of *I. attavipachi* (Fig. 2) is robust, with a sauropod-like straight, craniocaudally flattened shaft, with no indication of the sigmoid curvature usually seen in prosauropods (including large forms such as *Riojasaurus*¹⁸, although some melanorosaurids, such as *Camelotia*⁹, have fairly straight femora). There is a well defined, dorsomedially oriented articular head, unlike the condition in prosauropods, which have a more hook-shaped articular head, even in large plateosaurids¹⁹ and melanorosaurids^{9,10}. The greater trochanter is massive and bulging. There is no evidence of a lesser trochanter, unlike the condition in prosauropods⁸ and *Vulcanodon*¹. The fourth trochanter is in a very proximal position, as in some primitive dinosaurs²⁰. It forms a prominent S-shaped ridge on the caudal face of the shaft, close to the medial edge, ending distally in a slightly hook-shaped acute tip reminiscent of *Barapasaurus* and *Vulcanodon*. Although not wing-shaped as in prosauropods, the fourth trochanter of *Isanosaurus* is more prominent than in *Vulcanodon*, *Shunosaurus* and *Barapasaurus*. Its very peculiar shape, unlike the condition in other sauropods, is

considered as an autapomorphic character of this taxon, which otherwise mainly shows features that are plesiomorphic for sauropods. The strongly expanded distal end of the femur shows massive condyles, a well developed ectepicondyle (not usually seen in

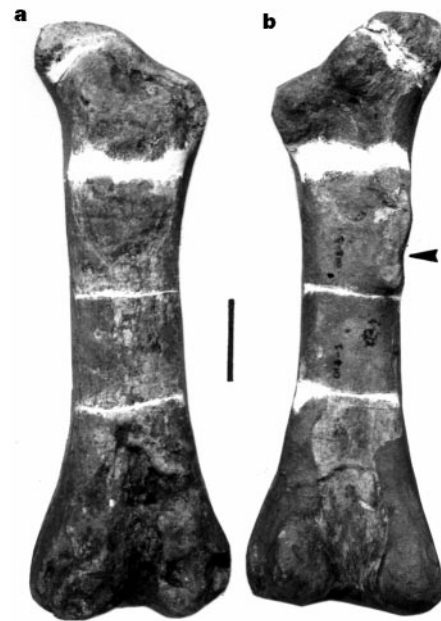


Figure 2 Left femur of *Isanosaurus attavipachi* (CH4-1). **a**, Anterior and **b**, posterior views. The arrow shows the peculiar S-shaped fourth trochanter. The white area between the fourth trochanter and the proximal articular head corresponds to a section of the specimen where the outer part of the bone is damaged but an inner continuity is present. Scale bar: 10 cm.

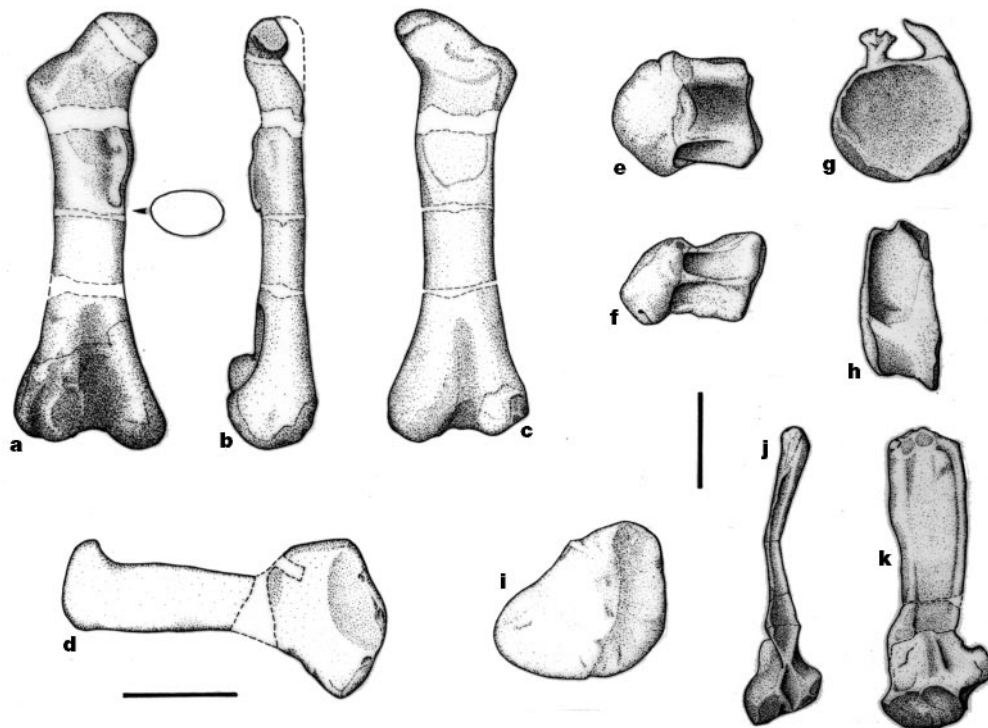


Figure 1 *Isanosaurus attavipachi*, elements of the holotype, palaeontological collection of the Department of Mineral Resources, Thailand, no. CH4. **a–c**, Left femur (CH4-1) in posterior (**a**, with cross-section at level of distal end of fourth trochanter), medial (**b**) and anterior (**c**) views. **d**, Right scapula in lateral view. **e, f**, Centrum of cervical vertebra (CH4-3) in left lateral (**e**) and ventral (**f**) views. **g, h**, Centrum of posterior dorsal vertebra

(CH4-6) in posterior (**g**) and left lateral (**h**) views. **i**, Right sternal plate in external view. **j, k**, Neural arch of posterior dorsal vertebra (CH4-7) in posterior (**j**), and right lateral (**k**) views. Horizontal scale bar (for **a–d**): 20 cm. Vertical scale bar (for **e–k**): 10 cm. Drawings by H.T.

prosaupods), and no longitudinal crest proximal to the lateral condyle (unlike prosaupods¹⁰).

I. attavipachi can clearly be placed among the Sauropoda because of the above-mentioned derived sauropod characters of its vertebrae and femur, which separate it from the Prosaupoda. Its primitive features are not particularly reminiscent of the Prosaupoda; rather, they seem to illustrate an early stage in the evolution of characters more fully developed in later, more advanced sauropods. Comparisons with other primitive sauropods reveal differences (notably in the femur), but their phylogenetic significance is uncertain. Although there is no consensus about the relationships of the oldest sauropods, recent phylogenies^{15,16,21,22} consistently place *Vulcanodon* in a very basal position; *Gongxianosaurus* also exhibits a number of primitive features reminiscent of prosaupods¹³. Comparisons between *Isanosaurus* and *Vulcanodon* are difficult, because few significant elements are known in both, although their femora are different. The opisthocelous cervical vertebrae of *Isanosaurus* show that it is more advanced than *Gongxianosaurus*, in which there are no opisthocelous vertebrae¹³; their femora also appear to be different. Uncertainties about the interrelationships of early sauropods, as expressed by the common use of the paraphyletic family Vulcanodontidae¹⁵, make it difficult to assess the exact phylogenetic and systematic position of *Isanosaurus*. A detailed analysis of early sauropod phylogeny being outside the scope of this paper, we refer *Isanosaurus* to Sauropoda incertae sedis.

The discovery of *I. attavipachi* not only shows that by late Triassic times the Sauropoda had already appeared, but also suggests that they must have had a relatively long and almost completely unknown evolutionary history in the Late Triassic, during which they coexisted with another group of large-bodied, heavily built sauropodomorphs, the melanorosaurid prosaupods. This is not unexpected, as calibrated phylogenies of the Sauropoda^{16,22} all show the history of the group extending well down into the Late Triassic. However, this assumption was theoretical and based mainly on the idea that the Sauropoda are the sister-group of the Prosaupoda. The remains of *I. attavipachi* are the first osteological evidence demonstrating the existence of Triassic sauropods. Previously, the only tentative fossil evidence for Triassic sauropods consisted of footprints^{4,5} (especially *Deuterosauropodopus*²³, from Lesotho), the attribution of which to sauropods is controversial^{4,5,24,25}.

Northeastern Thailand was already linked to China in the Late Triassic²⁶, and the earliest well attested sauropod is thus an Asian form. Even if ichnological evidence from the Late Triassic of southern Africa is inconclusive, *Vulcanodon* definitely indicates that sauropods occurred there at the very beginning of the Jurassic. Convincing sauropod footprints have been reported from the basal Jurassic (Hettangian) of Poland²⁷ and Italy²⁸. All this suggests that by the time of the Triassic–Jurassic boundary, sauropods already had a vast geographical distribution, doubtless made possible by Pangaeon palaeogeographical conditions. □

Received 2 November 1999; accepted 18 May 2000.

- Cooper, M. R. A reassessment of *Vulcanodon karibaensis* Raath (Dinosauria: Saurischia) and the origin of the Sauropoda. *Palaeontol. Afr.* **25**, 203–231 (1984).
- Jain, S. L., Kuttu, T. S., Roychowdhury, T. & Chatterjee, S. in *IV International Gondwana Symposium* 204–216 (Hindustan Publishing Corporation, Delhi, 1979).
- Yadagiri, P. A new sauropod *Kotasaurus yamanpalliensis* from Lower Jurassic Kota Formation of India. *Rec. Geol. Surv. India* **11**, 102–127 (1988).
- Ellenberger, F. & Ellenberger, P. Principaux types de pistes de vertébrés dans les couches du Stormberg au Basutoland (Afrique du Sud). *C. R. Somm. Séanc. Soc. Géol. Fr.* **4**, 65–67 (1958).
- Charig, A. J., Attridge, J. & Crompton, A. W. On the origin of the sauropods and the classification of the Saurischia. *Proc. Linn. Soc. Lond.* **176**, 197–221 (1965).
- Racey, A. et al. Stratigraphy and reservoir potential of the Mesozoic Khorat Group, NE Thailand. *J. Petrol. Geol.* **19**, 5–40 (1996).
- Buffetaut, E., Martin, V., Sattayarak, N. & Suteethorn, V. The oldest known dinosaur from southeast Asia: a prosauropod from the Nam Phong Formation (late Triassic) of northeastern Thailand. *Geol. Mag.* **132**, 739–742 (1995).
- Galton, P. M. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 320–344 (Univ. California Press, Berkeley, 1990).
- Galton, P. M. Notes on the Melanorosauridae, a family of large prosauropod dinosaurs (Saurischia: Sauropodomorpha). *Geobios* **18**, 671–676 (1985).

- Gauffre, F. X. The most recent melanorosaurid (Saurischia, Prosaupoda), Lower Jurassic of Lesotho, with remarks on the prosauropod phylogeny. *N. Jb. Geol. Paläont. Mh* **11**, 648–654 (1993).
- Barrett, P. A sauropod dinosaur from the Lower Lufeng Formation (Lower Jurassic) of Yunnan Province, People's Republic of China. *J. Vert. Paleontol.* **19**, 785–787 (1999).
- Dong, Z., Zhou, S. & Zhang, Y. The dinosaurian remains from Sichuan Basin, China. *Palaeontol. Sinica C* **23**, 1–145 (1983).
- He, X., Wang, C., Liu, S., Zhou, F., Liu, T. Cai, K. & Dai, B. A new species of sauropod from the Early Jurassic of Gongxian County, Sichuan. *Acta Geologica Sichuan* **18**, 1–7 (1998).
- Zhang, Y. *The Middle Jurassic Dinosaur Fauna from Dashanpu, Zigong, Sichuan* vol. 1 1–89 (Sichuan Publishing House of Science and Technology, Chengdu, 1988).
- Upchurch, P. The phylogenetic relationships of sauropod dinosaurs. *Zool. J. Linn. Soc.* **124**, 43–103 (1998).
- Wilson, J. A. & Sereno, P. C. Early evolution and higher-level phylogeny of sauropod dinosaurs. *Soc. Vert. Paleontol. Mem.* **5**, 1–68 (1998).
- Bonaparte, J. F. Evolucion de las vertébras presacras en Sauropodomorpha. *Ameghiniana* **36**, 115–187 (1999).
- Bonaparte, J. F. Los tetrapodos del sector superior de la formación Los Colorados; La Rioja, Argentina (Triásico superior). *Opera Lilloana* **22**, 1–183 (1971).
- Wellnhofer, P. Prosaupod dinosaurs from the Feuerletten (Middle Norian) of Ellingen near Weissenburg in Bavaria. *Rev. Paléobiologie* vol. **spec. 7**, 263–271 (1993).
- Sues, H. D. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 143–147 (Univ. California Press, Berkeley, 1990).
- McIntosh, J. S. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 345–401 (Univ. California Press, Berkeley, 1990).
- Upchurch, P. The evolutionary history of sauropod dinosaurs. *Trans. R. Soc. Lond. B* **349**, 365–390 (1995).
- Ellenberger, P. Contribution à la classification des pistes de vertébrés du Trias. Les types du Stormberg d'Afrique du Sud (I). *Palaeovertebrata Mém. Extr.* 1–104 (1972).
- Haubold, H. *Saurierfahrten* (A. Ziemsen, Wittenberg-Lutherstadt, 1984).
- Thulborn, T. *Dinosaur Tracks* (Chapman and Hall, London, 1990).
- Buffetaut, E. & Suteethorn, V. in *Biogeography and Geological Evolution of SE Asia* (eds Hall, R. & Holloway, J. D.) 83–90 (Backhuys, Leiden, 1998).
- Gierlinski, R. & Sawicki, R. New sauropod tracks from the Lower Jurassic of Poland. *Geol. Quart.* **42**, 477–480 (1998).
- Avanzini, M., Frisia, S., Van den Driessche, K. & Keppens, E. A dinosaur tracksite in an early Liassic tidal flat in northern Italy: paleoenvironmental reconstruction from sedimentology and geochemistry. *Palaios* **12**, 538–551 (1997).

Acknowledgements

This work was supported by the Geological Survey of Thailand, a special grant from the Institut National des Sciences de l'Univers (Paris), the Association Dinosauria (Espéraza, France), the Jurassic Foundation (Drumheller, Canada) and a NERC grant to G.C. We thank A. Milner (The Natural History Museum, London) for access to specimens in her care, and G. Olshevsky and T. Ford for bibliographic help.

Correspondence and requests for materials should be addressed to E.B. (e-mail: Eric.Buffetaut@wanadoo.fr).

Biochemical evidence of cannibalism at a prehistoric Puebloan site in southwestern Colorado

Richard A. Marlar[†], Banks L. Leonard[‡], Brian R. Billman[§], Patricia M. Lambert^{||} & Jennifer E. Marlar[†]

[†] Department of Pathology, University of Colorado School of Medicine, Denver, Colorado 80262, USA

[‡] Colorado Archaeological Society, Denver, Colorado 80250, USA

[§] Soil Systems Inc., Phoenix, Arizona 85004, USA

^{||} Department of Anthropology, University of North Carolina-Chapel Hill, Chapel Hill, North Carolina 27599, USA

^{||} Department of Sociology, Social Work and Anthropology, Utah State University, Logan, Utah 84322, USA

The existence of cannibalism is one of the most controversial issues in the archaeology of the American Southwest. Disarticulated, cut-marked and heat-altered human remains from non-burial contexts at prehistoric Puebloan (Anasazi) archaeological sites in the Four Corners region of the American Southwest have been interpreted by some scholars as evidence of cannibalism¹. Osteological studies indicate that many of the disarticulated

bodies found at these sites were processed in a manner consistent with food preparation². Opponents of this interpretation point out that non-cannibalistic practices such as secondary interment, corpse mutilation and ritualized witch executions might account for the assemblages³⁻⁷. Osteological evidence alone does not document the actual ingestion of human flesh. Here we show consumption of human flesh did occur as demonstrated in preserved human waste containing identifiable human tissue remains from a site with osteological evidence of cannibalism.

Sometime around AD 1150 a small Puebloan habitation site (5MT10010) located along Cowboy Wash in southwestern Colorado was suddenly abandoned^{8,9}. The site inhabitants' principal residences were three pithouses (Features 3, 13 and 15; Fig. 1). Several lines of evidence indicate that during the abandonment or soon after, the bodies of seven people of both sexes and various ages were disarticulated, defleshed and apparently cooked as if for consumption by other humans⁹⁻¹². Their incomplete remains were left directly on floors and in other non-burial contexts in two of the pithouses (Features 3 and 13; Fig. 2)^{8,9}.

The contexts and types of artefacts left behind in the pithouses and the conditions of their roofs indicate that the pithouses at 5MT10010 were suddenly abandoned^{8,9}. This site was excavated as part of a larger archaeological study of 17 Puebloan sites on the

southern piedmont of Sleeping Ute Mountain. The project involved the excavation of 105 structures, including 36 pithouses or pitstructures dating from AD 450-1280 (refs 9, 10). The abandonment observed in the pithouses at 5MT10010 differed markedly from the pattern seen at the other sites excavated during the project. The typical pattern of structure abandonment involved removal of virtually all artefacts and materials of value. Grinding stones, finely polished tools, ornaments and whole vessels were rarely left behind. Structural wood and stone, especially shaped slabs, were routinely scavenged for re-use. In cases where roofing materials were not stripped, the roof was typically set ablaze after useable artefacts had been removed from the structure.

In contrast, at 5MT10010, household goods, such as cooking pots and serving wares, valuable items, such as ornaments and polished stone tools, and salvageable construction materials, such as shaped stone slabs and wooden posts, were left in place in all the pithouses at abandonment. Many of the vessels, tools, ornaments and shaped stones were found directly on floor and bench surfaces with no sediment underneath, indicating that they were found at or near where they were originally left. Microstratigraphic evidence also indicates that the roofs of all the pithouses decayed gradually in place, rather than being burned or scavenged for re-use as was done with virtually all other southern Piedmont habitation structures^{8,9}.

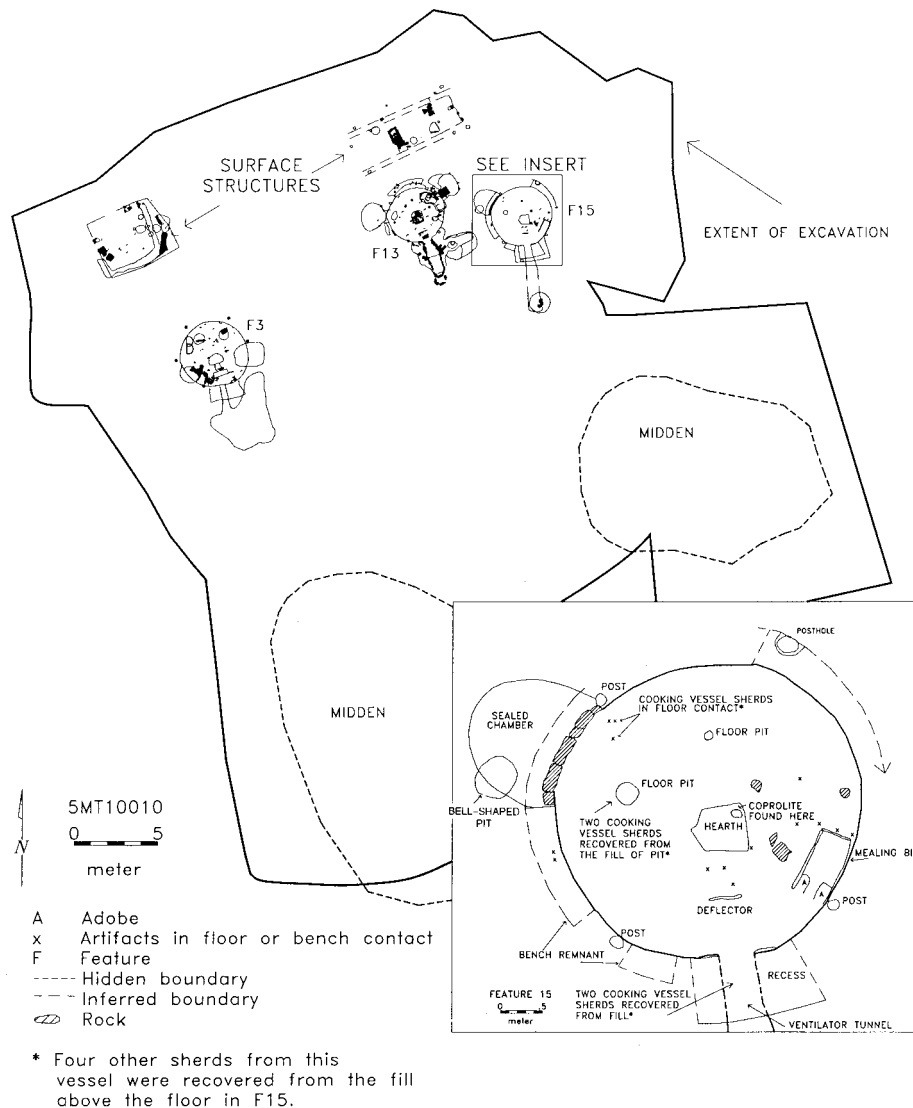


Figure 1 Site map of 5MT10010, showing the three residential pithouses (Features 3, 13 and 15; F3, F13 and F15) with the associated surface structures and trash middens that were in use when the site was abandoned. The inset depicts the interior of Feature 15,

showing where the coprolite was recovered from the hearth and where cooking pot shards were recovered from the structure. The location of four shards recovered from fill above the floor is not indicated.

The disarticulated human remains were found scattered and piled in similar contexts to the valuable artefacts^{8–12}. In Feature 3, over a thousand human bones and fragments were found piled in a side chamber while others were recovered directly from the floor of the structure, with no sediment underneath (Fig. 1). The remains represented a minimum of four adults and one adolescent. In Feature 13, whole bones and fragments were left directly on the floor, piled in a side chamber and stacked on a bench. Scorched tooth and bone fragments were also found in the central hearth and in ash piles on the structure floor. The bones in Feature 13 were from two subadults.

Other things were left in the pithouses during or soon after the site was abandoned^{8,9}. A set of stone tools consistent with use in butchering was scattered around the hearth on the floor of Feature 13. Several of the tools were tested by crossover immunoelectrophoresis for blood residues; two cutting tools tested positive for human blood^{13,14}. Although no human remains were left behind in the third pithouse at the site (Feature 15) near the time of abandonment, fragments of a cooking pot were found scattered throughout the structure. Some of the fragments were in direct contact with the floor (Fig. 1, inset). Finally, an unburned human faecal deposit (coprolite) was found in the ashy fill of the structure hearth (Fig. 1, inset). Its unburned condition demonstrated that it was deposited after the last use of the hearth. This was the only coprolite recovered from the site and may be the only one identified from a structure hearth from anywhere in the American Southwest.

The abandonment of a cooking pot in Feature 15 opened the possibility that biochemical analyses might detect human tissue residues, supporting the hypothesis that human body parts were cooked. An immunological detection assay method (ELISA) has been used to identify animal meat residues in cooking pots from archaeological contexts¹⁵. To test for the cooking of human muscle tissue in ceramic vessels, 11 shards from the Feature 15 cooking vessel were analysed for human myoglobin. Myoglobin is a protein molecule that transports oxygen from the inner surface of the membrane of skeletal and cardiac muscle cells to the energy-generating components within the cells. Five shards from other vessel types, or from vessels that were already broken before the events surrounding the abandonment of 5MT10010 began, were also analysed for human myoglobin. One of these was from the floor of Feature 3, where it was found lying directly under the face of a

disarticulated human adolescent. The other four were from the floor of Feature 13, the same structure where blood residues were detected on cutting tools. Only the shards from the cooking vessel in Feature 15 tested positive for human myoglobin (2.8–48 µg of human myoglobin per shard).

For controls, 29 shards from other archaeological sites were tested using the same procedures: 14 cooking vessel shards from a midden area associated with a contemporaneous Pueblo II/Pueblo III (AD 1075–1175) site (5MT5501) from southwestern Colorado, and 15 shards from an intermittent campsite (5JF321) southwest of Denver that contained a Woodland Ceramic Tradition component (AD 150–1150) with associated shards from a minimum of 6–8 cooking vessels¹⁵. All control shards were negative for human myoglobin (< 1 ng per sample). The presence of human myoglobin only on cooking vessel shards from 5MT10010 is consistent with the hypothesis that human muscle tissue was cooked in that vessel.

The discovery of a coprolite that was deposited near the time of abandonment of the site, during or shortly after butchering and cooking of human remains, provided the potential to yield direct evidence of cannibalism. The coprolite was found in Feature 15 (Fig. 1) and consisted of a single mass of desiccated faecal material, 30 g dry weight, of a size and shape consistent with human origin^{14,16}. The position and condition of the coprolite indicated that it was defecated directly into the cold hearth in Feature 15 (Fig. 1, inset)^{8,9}. Macroscopic analysis of the human coprolite revealed no detectable plant remains, which is extremely unusual for an ancient Puebloan coprolite^{14,16–18}. Microscopic analysis indicated that starch granules and phytoliths were virtually absent. The absence of starch granules is considered a strong indicator that maize in particular was not part of the meal(s) represented in the coprolite¹⁹. The only pollens identified were from *Cheno-Am*, low-spine *Compositae*, and trace amounts from *Poaceae*, all of which could have derived from wind-borne, ambient pollen¹⁶. The absence of plant remains except for these pollen types is consistent with the hypothesis that the depositor of the coprolite had not consumed plant foods 12–36 h before defecation. Although bone fragments and keratinous elements, such as hair, were not detected among the gross contents, the absence of plant material indicated that the meal(s) represented by the coprolite were probably composed entirely of meat^{14,16}.

To test the hypothesis that human flesh was consumed, it was

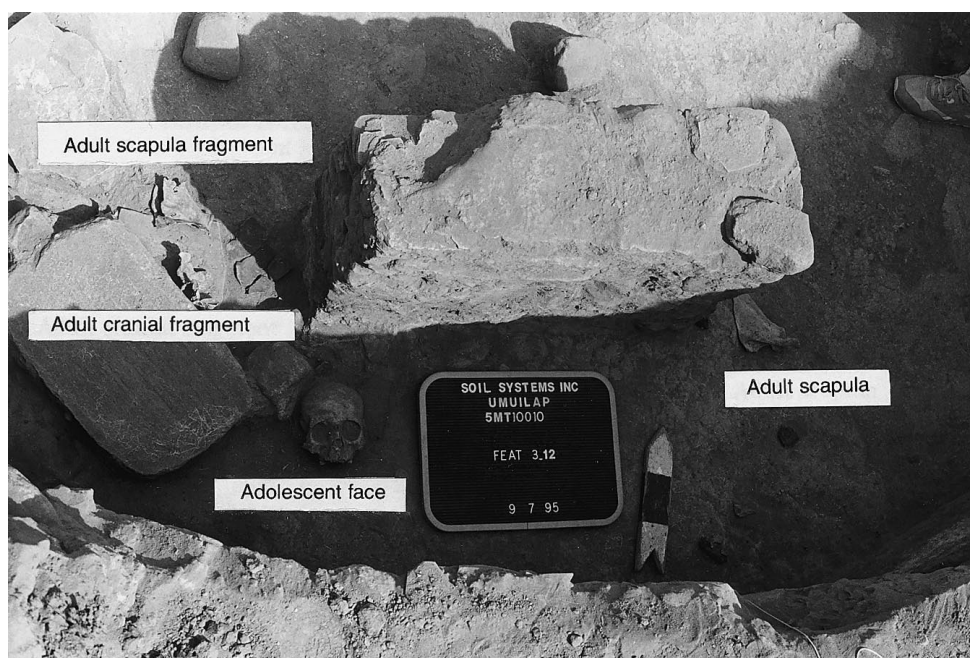


Figure 2 Human bones on the floor of Feature 3, south of the structure hearth.

necessary to identify a human-derived substance in the coprolite², but many human molecules normally occur in human stool material. For example, cells from the intestinal lining are constantly shed during the peristaltic process and blood from intestinal lesions may be present in stool samples. Therefore, it was necessary to identify a human substance that could only be present in the coprolite because it had been consumed by the depositor and could not be derived from his/her own tissues during digestion and elimination. Myoglobin is found only in skeletal and cardiac muscle cells, and not found in cells of the blood, skin, connective tissue, vascular tissue, tissues of the lymphatic system, nor in the smooth muscle cells of the digestive system. Therefore, human myoglobin should only be present in faecal material if it is consumed and passed through the digestive system by the depositor of the faeces. Furthermore, the chemical composition of myoglobin differs among animal taxa, making it possible to identify the type (species) of flesh consumed (J.E.M., unpublished results). Consequently, the ELISA technique can distinguish the presence of taxon-specific meat remains in the faeces of meat consumers. Bovine myoglobin, for example, was detected in samples from modern individuals that had consumed cooked beef within the last 24 h, demonstrating that taxon-specific myoglobin can be detected in faecal material.

Analysis of the coprolite from Feature 15 by ELISA detected human myoglobin (18–62 ng of human myoglobin per g of coprolite). The amount of myoglobin (> 5 s.d. above the average of the negative control) was lower than the amount detected on some of the shards from Feature 15 (7–10 s.d.). Apparently, the majority of the human myoglobin was broken down (degradation and hydrolysis) in the cooking process and in the gastrointestinal system of the consumer, and only a small amount remained in the coprolite that was recognizable to the human myoglobin-specific purified antibody. Human myoglobin was undetectable (< 5 s.d.) in 39 modern human faecal extracts used as controls, including samples from patients with positive blood in the stool sample. Furthermore, 20 prehistoric coprolites were tested as controls and showed no human myoglobin (< 5 ng of human myoglobin per gm of coprolite). The control coprolites were from Salmon Ruin, an open-air Puebloan site with occupation contemporaneous to 5MT10010. Although a possible cannibalism assemblage has been described from Salmon Ruin¹, all of the control coprolites were recovered from a deep

latrine deposit that clearly predates events surrounding the formation of the possible cannibalism assemblage (K. Reinhard, personal communication). To rule out contamination from insects in the coprolite from 5MT10010, internal larval proteins were tested for crossreactivity with human myoglobin; the results of these tests were negative.

Direct evidence for the consumption of human tissue by humans is necessary to demonstrate definitively that human cannibalism occurred at an archaeological site. Previous archaeological and osteological studies have strongly indicated that cannibalistic episodes took place among the ancient Pueblos, but the evidence has been essentially circumstantial. The analysis of the coprolite and associated remains from 5MT10010 at last provides definitive evidence for an episode of cannibalism involving ancient Pueblos. Results of the human myoglobin ELISA analyses of the human coprolite and shards from a ceramic vessel are consistent with the archaeological and osteological evidence of cannibalism at 5MT10010. During or after the sudden abandonment of the site, disarticulated, defleshed and heat-altered human remains were left in non-burial contexts in association with butchering tools with human blood residue, a cooking vessel with human myoglobin residue and a human coprolite containing human myoglobin. These data demonstrate that humans both processed and consumed human flesh at the site.

Cannibalism has occurred in a wide range of societies for a wide variety of reasons, including starvation, ancestor worship and political terrorism^{20–23}. With the presentation of the first direct evidence of cannibalism in the American Southwest in the prehistoric era, we hope that the debate will shift from the question of whether or not cannibalism occurred to questions concerning the social context, causes and consequences of these events. □

Methods

Artefact, coprolite and stool sample processing

We processed the shards, coprolite and control samples in an identical manner. We immersed the shards in artefact buffer (0.02 M Tris, 0.5 M NaCl, 0.5% Triton X-100, pH 7.4), sonicated them for 2 h and centrifuged them to remove particulate matter. We removed Triton X-100 by dilution/concentration three times using ultra-filtration membranes (cut-off at relative molecular mass < 10,000; Amicon). We dissolved the coprolites (100 mg) and control stool samples (500 mg) in artefact buffer and processed them as for the shards. The final volume was one-fifth the starting volume.

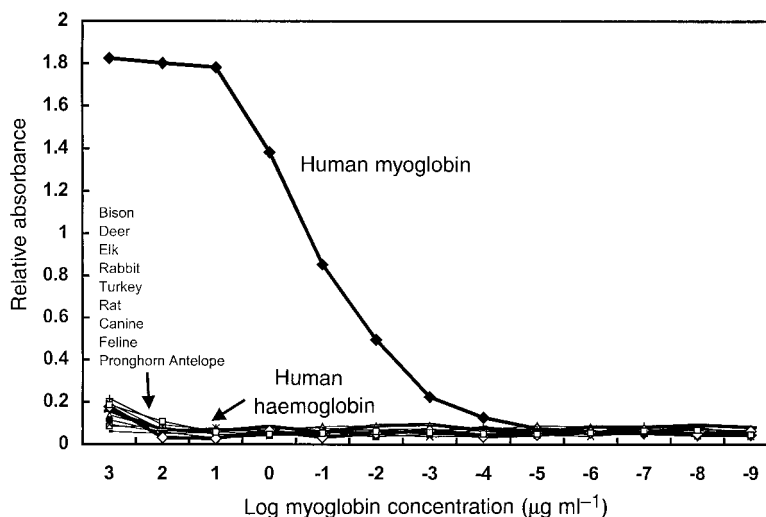


Figure 3 Specificity and sensitivity of the myoglobin assay. Dose–response curves from the ELISA assay demonstrate the specificity of the immuno-purified human myoglobin antibody toward human myoglobin, human haemoglobin and myoglobin from other animal species. The concentrations of myoglobin from each species were determined using a commercial protein assay. The myoglobin samples were assayed by

serial log-dilution for each myoglobin/haemoglobin sample. Each species of myoglobin was purified (> 90% myoglobin) by ion-exchange chromatography after extraction from skeletal muscle tissue. The other myoglobin species tested were bison, deer, elk, rabbit, turkey, rat, canine, feline and pronghorn antelope. Purified human haemoglobin was also tested.

Myoglobin detection assay

We used a sandwich-type ELISA to analyse for human myoglobin on shards, human coprolite samples and human stool samples. We applied a 100 µl aliquot of 1/1000 dilution of the capture antibody (immuno-purified rabbit anti-human myoglobin antibody from the purified immunoglobulin fraction, Sigma) in 0.05 M carbonate buffer, pH 9.6, to the plate overnight at 4 °C. We removed the unbound antibody by washing five times with ELISA wash buffer (0.025 M Tris, 0.14 M NaCl 0.025% Tween, pH 7.4) in an automated ELISA washer.

We diluted the sample and controls 1/100 in ELISA dilution buffer (0.5 M Tris, 0.14 M NaCl, 0.03 M KCl, 0.2% Tween, 0.4% PEG-8000, pH 7.4) and applied 100 µl to the appropriate wells for 1 h at 22 °C. After washing the wells (as above), we applied mouse monoclonal anti-human myoglobin (Sigma; 100 µl diluted 1/4,000 in ELISA dilution buffer) for 1 h at 22 °C. We washed the wells three times and applied the detection antibody (Sigma; 100 µl of sheep anti-mouse IgG conjugated to horse radish peroxidase, diluted 1/10,000 in ELISA dilution buffer) 1 h at 22 °C. We washed the wells three times and added the substrate (TMB/Urea; Sigma) for 5 min. We stopped the reaction with 2 M H₂SO₄ and read the plate at 450 nm on an ELISA reader (Dynex MRX, Chantilly, VA). We assayed each sample or control using six replicates, three times each by two individuals. We averaged the values from each experiment and compared them statistically to the negative controls using the Student's *t*-test. We considered the results as positive when *P* < 0.001 and at least 5 s.d. above the average negative control.

The commercial rabbit anti-human myoglobin antibody reacted minimally with myoglobin from several other species used as possible food sources. To remove these crossreacting antibodies, the rabbit anti-human myoglobin antibodies were immuno-adsorbed with different species of myoglobin (deer, bovine, sheep, antelope, rabbit, turkey, chicken, elk, mouse and rat). The individual myoglobin samples were coupled to Sepharose (Pharmacia) to bind the antibodies specific for the different species of myoglobin. The remaining human-specific antibodies were concentrated and used in the ELISA procedure. The immuno-purified polyclonal rabbit anti-human myoglobin antibodies recognized only human myoglobin in a dose-dependent manner (Fig. 3). The concentration of human myoglobin detected in the coprolite ranged from 18 to 62 ng ml⁻¹. No detectable concentrations of myoglobin were observed with serial dilutions of myoglobin (> 1 mg ml⁻¹) from the other species, including the 'food source' species found in the region (Fig. 3). Crossreactivity with non-human primates was not considered, because no evidence of non-human primates has been found in prehistoric archaeological contexts in the continental United States. Furthermore, the nearest contemporaneous non-human primate populations were located in tropical Mexico.

Artefact and faecal controls

The control shards from 5MT5501 were provided by Jerry Fetterman, Woods Canyon Archaeological Consultants, Inc. The site occupation was contemporaneous to the Cowboy Wash site (5MT10010), but lacked any indication of possible cannibalism. 5MT5501 is located about two miles west of Dolores and 18 miles north of Sleeping Ute Mountain, in southwestern Colorado. The control shards from 5JF321 were provided by the Colorado Archaeological Society from their excavation about ten miles southwest of Denver in the Ken Caryl Valley¹⁵. These shards are of the Woodland Ceramic Tradition. Some control shards from both 5MT5501 and 5JF321 were positive for deer and rabbit myoglobin and/or blood, but control shards from ancient Pueblo or Plains cultures did not contain human myoglobin residue.

Control faecal tests were conducted to determine whether human myoglobin was present in faeces from modern normal individuals (25 samples), modern individuals with blood in their stool samples (ten samples), or modern individuals who had consumed cooked beef within 24 h of defecation of the specimen (four samples). These controls did not show detectable levels of human myoglobin (< 5 s.d. of the average negative control). This result is consistent with the hypothesis that human myoglobin is not derived from the tissues of a defecator, even when the stool sample is positive for blood. In contrast, the control samples from the beef consumers tested positive for bovine myoglobin, demonstrating that orally ingested myoglobin can survive the processes of cooking and digestion, can be detected in human faecal material, and can be identified as to biological taxon of origin. The modern stool samples were collected for clinical testing and the remaining material was considered 'discarded specimen material' from the clinical laboratory. The only personal information available to the authors was the patient's occult blood status.

Received 7 March; accepted 6 June 2000.

1. Turner, C. G. & Turner, J. A. *Man Corn: Cannibalism and Violence in the American Southwest* (Univ. of Utah Press, Salt Lake City, Utah, 1999).
2. White, T. D. *Prehistoric Cannibalism at Mancos 5MTUMR-2346* (Princeton Univ. Press, Princeton, New Jersey, 1992).
3. Baker, S. *Rattlesnake Ruin (42SA1853): A Case of Violent Death and Perimortem Mutilation in Anasazi Culture of San Juan County, Utah*, Thesis, Brigham Young Univ., Provo Utah (1990).
4. Bullock, P. A reappraisal of Anasazi cannibalism. *Kiva* 57, 5–16 (1991).
5. Darling, J. A. Mass inhumation and the execution of witches in the American Southwest. *Am. Anthropol.* 100, 732–752 (1999).
6. Nass, G. G. & Bellatoni, N. F. A prehistoric multiple burial from Monument Valley evidencing trauma and possible cannibalism. *Kiva* 47, 257–271 (1982).
7. Dongoske, K. E., Martin, D. L. & Ferguson, T. J. Critique of the claim of cannibalism at Cowboy Wash. *Am. Antiquity* 65, 179–190 (2000).
8. Leonard, B. L., McAndrews, K. & Lambert, P. M. in *The Puebloan Occupation of the Ute Mountain Piedmont* Vol. 3 (ed. Billman, B.) Ch. 3 (Soil Systems Publications in Archaeology No. 22, Phoenix, in the press).
9. Billman, B. R., Lambert, P. M. & Leonard, B. L. Cannibalism, warfare, and drought in the Mesa Verde region during the twelfth century A. D. *Am. Antiquity* 65, 145–178 (2000).

10. Billman, B. R. *The Puebloan Occupation of the Ute Mountain Piedmont: Synthesis and Conclusions*, Vol. 7 (Soil Systems Publications in Archaeology No. 22, Phoenix, in the press).
11. Lambert, P. M. in *The Puebloan Occupation of the Ute Mountain Piedmont* Vol. 5 (ed. Billman, B.) Ch. 6 (Soil Systems Publications in Archaeology No. 22, Phoenix, 1999).
12. Lambert, P. M., Billman, B. R., & Leonard, B. L. Explaining variability in mutilated human bone assemblages from the American Southwest: A case study from the Southern Piedmont of Sleeping Ute Mountain, Colorado. *Int. J. Osteoarchaeol.* 10, 49–64 (2000).
13. Newman, M. E. in *The Puebloan Occupation of the Ute Mountain Piedmont* Vol. 3 (ed. Billman, B.) Appendix A (Soil Systems Publications in Archaeology No. 22, Phoenix, in the press).
14. Lambert, P. M. *et al.* Response to critique of the claim of cannibalism at Cowboy Wash. *Am. Antiquity* 65, 397–406 (2000).
15. Hammond, W. & Rathbun, F. Preliminary report of the 1998 excavations at Swallow site (5JF321>). Denver Chapter, Colorado Archaeological Society Report, 22–24 (1998).
16. Reinhard, K. J. in *The Puebloan Occupation of the Ute Mountain Piedmont*, Vol. 3 (ed. Billman, B.) Appendix B (Soil Systems Publications in Archaeology No. 22, Phoenix, in the press).
17. Minnis, P. E. Prehistoric diet in the Southwest: Macroplant remains from Four Corners feces. *Am. Antiquity* 54, 543–563 (1989).
18. Reinhard, K. in *Diet, Demography and Disease: Changing Perspectives of Anemia* (eds Stuart-McAdam, P. & Kent, S.) 219–258 (Aldine de Gruyter, New York, 1992).
19. Vinton, S. D. *Dietary Analysis of Coprolites from Northern Chile*. Thesis, Univ. of Nebraska (1997).
20. Brown, P. & Tuzin, D. F. *The Ethnography of Cannibalism* (Soc. Psych. Anthropol. Publishers, Washington DC, 1983).
21. Chagnon, N. A. *Yanonomamö* (Harcourt, Brace, Jovanovich College, Fort Worth, Texas, 1997).
22. Grayson, D. K. Donner party deaths: A demographic assessment. *J. Anthropol. Res.* 46, 223–242 (1990).
23. Keenleyside, A., Bertulli, M. & Fricke, H. C. The final days of the Franklin Expedition: New skeletal evidence. *Arctic* 50, 36–46 (1997).

Acknowledgements

We thank T. G. Knight (Ute Mountain Ute Tribe), W. Hurley (Bureau of Reclamation), B. Harrill (Bureau of Indian Affairs), C. D. Breternitz (President of Soil Systems, Inc. and Principal Investigator) and C. K. Robinson (Projects Manager of Soil Systems, Inc.) for support of the excavation of 5MT10010; D. C. Kressin, A. Marlar and A. Schultz for help with the human myoglobin assay; the field crew at 5MT10010; and J. Merewether and M. Martin (Laboratory Directors for Soil Systems, Inc.). C. Murphy originally encountered the coprolite and E. Hansen first identified it as a human faecal deposit. R.M. thanks W. Hammond for encouraging the establishment of a biomolecular archaeology laboratory.

Correspondence and requests for materials should be addressed to R.A.M. (e-mail: marlarr@den-res.org).

Uptake of dissolved organic carbon and trace elements by zebra mussels

Hudson A. Roditi, Nicholas S. Fisher & Sergio A. Sañudo-Wilhelmy

Marine Sciences Research Center, State University of New York, Stony Brook, New York 11794-5000, USA

Zebra mussels (*Dreissena polymorpha*) are widespread and abundant in major freshwater ecosystems in North America, even though the phytoplankton food resources in some of these systems seem to be too low to sustain them^{1,2}. Because phytoplankton biomass is greatly depleted in ecosystems with large *D. polymorpha* populations^{3,4} and bacteria do not seem to be an important food source for this species⁵, exploitation of alternative carbon sources may explain the unexpected success of *D. polymorpha* in such environments. Here we examine the possibility that absorption of dissolved organic carbon (DOC) from water^{6–9} could provide a nutritional supplement to zebra mussels. We find that mussels absorb ¹⁴C-labelled DOC produced by cultured diatoms with an efficiency of 0.23%; this indicates that DOC in natural waters could contribute up to 50% of the carbon demand of zebra mussels. We also find that zebra mussels absorb some dissolved metals that have been complexed by the DOM; although absorption of dissolved selenium was unaffected by DOC, absorption of dissolved cadmium, silver and mercury by the mussels increased 32-, 8.7- and 3.6-fold, respectively, in the presence of high-molecular-weight DOC.

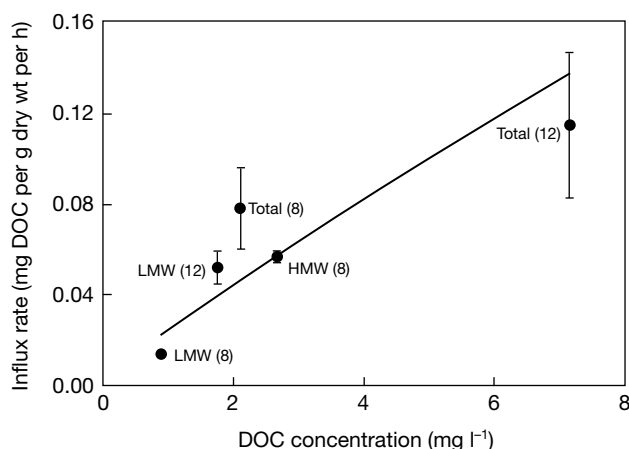


Figure 1 Influx of algal DOC into zebra mussels. Values shown are means $\pm$ 1 s.d.; the number of replicate mussels are shown in parentheses next to the data points. Values of a and b in the equation $I = aQ^b$ (see text) were calculated as the y -intercept and slope, respectively, of the regression line through log-transformed data. The regression was performed on log-transformed data to reduce heteroscedasticity and the influence on the regression slope of the point at 7.2 mg DOC l^{-1} . Regression equation: $\log \text{influx} = 0.872(\pm 0.077) \times \log(\text{DOC concentration}) - 1.61(\pm 0.038)$, $r^2 = 0.734$. The uptake rate constant a ranged from 0.023 to 0.0271 per g dry wt per d, with a resulting absorption efficiency of $0.23 \pm 0.02\%$.

Table 1 DOC influx rates into zebra mussels

Treatment	DOC (mg l^{-1})	DOC-normalized influx rate (mg DOC per g dry wt per h/mg DOC)
LMW	1.75	0.030 ± 0.004
LMW	0.88	0.016 ± 0.002
HMW	2.66	0.021 ± 0.001
Total	7.16	0.016 ± 0.004
Total	2.10	0.037 ± 0.009

These rates have been normalized to DOC concentration in exposure water. Values are mean $\pm$ 1 s.d.

Among the studies which have explored the nutritional aspects of DOC for freshwater invertebrates^{6,10}, significant exploitation of DOC was only found for some insect larvae, which can use mucopolysaccharide colloidal matter exuded from cells¹¹ as a nutritional source^{10,12}. Furthermore, many studies have shown that DOC complexes trace metals and reduces their bioavailability for a variety of aquatic organisms¹³, although exceptions have been noted in which the presence of DOC enhances metal influx rates^{13–15}. A better understanding of the factors, including DOC, that influence metal bioavailability is important for evaluating metal bioaccumulation and toxicity in aquatic environments. Here we describe experiments which measured absorption of radiolabelled DOC and associated trace metals by *D. polymorpha*.

Influx rates of algae-derived $DO^{14}C$ increased with ambient concentration and without consistent influence of DOC molecular size (Fig. 1, Table 1). Influx rate (I , mg per g dry wt mussel per h) can be described as $I = aQ^b$, where a is the uptake rate constant ($l g^{-1} h^{-1}$), Q is the solute concentration ($mg l^{-1}$), and b is the power of the relationship, and the uptake rate constant a is equal to the product of the absorption efficiency (%) and the filtration rate ($l g^{-1} h^{-1}$)¹⁶. The uptake rate constant can thus be divided by the

filtration rate to calculate an absorption efficiency, which is equal to the fraction of total carbon pumped that is absorbed by the mussel. Filtration rates estimated here from individually measured soft-tissue dry weights and published models¹⁷ were $10.9 \pm 0.2 l g^{-1} h^{-1}$ (mean $\pm$ s.d.), comparable with independent observations¹⁸ for a range of particle concentrations and types. Regression analysis produced an uptake rate constant of $0.0245 l g^{-1} h^{-1}$ with a resulting absorption efficiency of DOC of $0.23 \pm 0.02\%$. The value of b in the influx equation $I = aQ^b$ was 0.872, indicating that the increase in DOC influx with increasing ambient DOC concentration was slightly less than linear.

Calculations showed that if all water in zebra mussel tissues were to be exchanged with ambient water (and its DOC) in the timescale of these experiments, $\leq 14\%$ of the DOC absorbed by the mussels exposed for 2 h could be attributed to passive uptake. An alternative explanation for the ^{14}C uptake was considered—that the zebra mussels filtered bacteria that grew on the $DO^{14}C$. However, calculations showed that the mussels filtered a volume equivalent to the entire volume in the feeding chambers every 9.5 ± 1.1 min. At this clearance rate, the zebra mussels would remove bacterial populations before they had sufficient time to grow and convert significant amounts of $DO^{14}C$ into particulate organic carbon (POC) in the feeding chambers. Given the short incubations and the filtration of experimental water immediately before experiments, it is unlikely that DOC formed aggregates¹⁹ filterable by *D. polymorpha*; consequently mussel uptake of the organic matter represented absorption of dissolved matter rather than particle filtration, consistent with autoradiographic evidence for DOC absorption in marine mussels⁸.

Table 2 shows the effects on metal concentration factors (CFs) of exposing the mussels to water with three different treatments: dosed with high-molecular-weight (HMW) DOC, with low-molecular-weight (LMW) DOC, and with total DOC. The metal CFs were determined for each metal as ng metal per g dry wt animal divided by ng metal per g water. For HMW and total DOC treatments, the CFs varied over an order of magnitude between elements, with $Ag > Hg > Cd > Se$, reflecting the order of metal absorption efficiencies (atoms absorbed/atoms pumped) (means of 1.87% for Ag, 1.17% for Hg, 1.02% for Cd, 0.03% for Se) in zebra mussels¹⁹. CFs of Ag, Cd and Hg in zebra mussel soft tissues were strongly influenced by DOC content, but uptake of Se (added as selenite), which does not readily get complexed by ambient DOC²⁰, was not affected by DOC (Table 2). CFs of Cd, Ag and Hg were 57-times, 6.3-times, and 4.1-times higher, respectively, at 3.28 mg (DOC) l^{-1} (in the total DOC treatment) than at 1.96 mg (DOC) l^{-1} (in the LMW DOC treatment), indicating that Cd was most affected by the DOC. The HMW and total DOC treatments differed from the LMW treatment both quantitatively, having means of 1.3-times and 1.7-times more DOC, respectively, and qualitatively (molecular size). When normalized to DOC concentration to isolate molecular size of the DOC as a variable, CFs were higher in either total or HMW DOC than in LMW treatments ($P < 0.05$) for Cd (means 24–34 times greater), Ag (3.8–6.6 times), and Hg (2.4–2.7 times), indicating that metal uptake was enhanced by the presence of HMW DOC.

The extent to which the fresh DOC in our experiment represents naturally occurring DOC which undergoes ageing and microbial decomposition is unknown, although recent measurements of DOC composition in rivers in the northeastern United States do not differ greatly from freshly produced DOC in phytoplankton exudates

Table 2 DOC content and trace-element concentration factors

Treatment	DOC content	Concentration factors (ng metal per g dry wt/ng metal per ml)			
		Ag	Cd	Hg	Se
LMW	<1 kDa: 1.96 ± 0.08	207 ± 23	8.0 ± 1.3	172 ± 31	59 ± 9
HMW	>1 kDa: $2.57 \pm 0.08^*$	$1,796 \pm 625$	253 ± 40	616 ± 81	60 ± 19
Total	<0.2 μm : 3.28 ± 0.09	$1,309 \pm 293$	452 ± 47	701 ± 170	70 ± 13

These concentration factors were determined in zebra mussels after 2-h exposures in LMW, HMW and total DOC treatments (mean $\pm$ 1 s.d., $n = 7$).

* This treatment also contained 0.37 mg l^{-1} of DOC < 1 kDa, making the total DOC concentration in this treatment 2.94 mg l^{-1} .

(D. Repeta, personal communication). This is consistent with the observation (before the zebra mussel invasion) that 69% of the allochthonous DOC in the Hudson River is respired by bacteria²¹. Here we consider the implications of DOC uptake by zebra mussels in the tidal freshwater Hudson River, where zebra mussel carbon assimilation has been estimated at 120–150 g C m⁻² yr⁻¹ and phytoplankton production is estimated to be threefold lower¹. With a mean Hudson River DOC concentration of 3.74 mg C l⁻¹ (ref. 22), a DOC uptake rate constant in *D. polymorpha* of 0.0245 l per g dry wt per h (0.23% absorption efficiency), and assuming 18 h of pumping activity per d (ref. 23), zebra mussels would take up 1.65 mg C per g dry wt per d, or 0.37% of their weight of carbon per d, which is approximately half of the no-growth carbon requirements for similar sized mussels at the same temperature²⁴. This maintenance ration may represent a carbon intake minimum to avoid starvation. Assuming that absorbed DOC is assimilated into the mussel tissues, DOC uptake may contribute to success of zebra mussel populations in turbid environments such as the Hudson with low food quality, despite predictions that they should not thrive².

When DOC is absorbed by zebra mussels, metals complexed by this DOC may be co-transported into the mussels, where they can deposit in diverse internal tissues²⁵; this is in addition to any free metal ions which may be absorbed. If accumulation of DOC-complexed metal does occur, then our metal accumulation results, in which metal CFs were higher in waters containing HMW DOC, can be explained by the greater complexation of these metals by HMW DOC than by LMW DOC. Measurements in the freshwater region of the Hudson indicate that 80–90% of 'dissolved' Ag and Hg is associated with HMW DOC²⁶. □

Methods

First attempts at analysing DOC absorption by *D. polymorpha* measured concentrations of natural DOC in filtered Hudson River water during 4-h exposures to mussels. No significant changes in DOC concentrations were detectable using a Shimadzu TOC 5000 analyser, however the variability in the analyses and the possibility that mussels could have imparted DOC into the water complicated interpretations of these data. Consequently, to reduce analytical error and produce an unambiguous data set, a culture of the euryhaline diatom *Thalassiosira pseudonana* was uniformly radiolabelled²⁷ with 660 kBq of NaH¹⁴CO₃, after which cells were filtered onto 1.0-µm polycarbonate membranes and lysed in 500 ml of 0.2-µm-filtered Hudson River water in which naturally-occurring DOC was photo-oxidized by ultraviolet irradiation²⁸ (DOC < 0.1 mg (C) l⁻¹).

To evaluate the importance of the molecular size of DOC compounds for absorption in the mussels, we compared total DOC (< 0.2 µm) treatments with LMW (< 1 kDa) and HMW (1 kDa–0.2 µm) DOC fractions. Differences in HMW and LMW composition (LMW organic material is richer in organic N, total aldose sugars and amino acids than HMW DOC)²⁹ have been demonstrated, and a large fraction of metals associate with HMW material in fresh waters³⁰. The extent to which these compositional differences affect metal bioavailability in freshwater systems is unknown. Diatom lysate was filtered through 0.2-µm membranes; the total DOC treatment consisted of 0.2-µm-filtered lysate diluted with photo-oxidized Hudson River water. LMW and HMW DOC treatments were prepared by diluting another aliquot of 0.2-µm filtered lysate with photo-oxidized Hudson River water, then ultrafiltering (1-kDa Millipore Prep/Scale column). The unamended filtrate was used for the LMW DOC treatment and the HMW DOC treatment was produced by diluting the retentate with photo-oxidized Hudson River water. DOC concentrations were calculated using the specific activity of 15 Bq per µg algal carbon in culture (derived from a final algal carbon load of 22.8 mg and a total radioactivity associated with this biomass of 333 kBq). The final DOC concentration in experimental water batches ranged from 0.9 to 7.3 mg l⁻¹ (0.9 and 1.9 mg l⁻¹ for LMW, 2.9 mg l⁻¹ for HMW, and 2.1 and 7.3 mg l⁻¹ for total DOC). Water for all treatments was filtered through 0.2-µm polycarbonate membranes immediately before addition of mussels and kept still to minimize turbulent shear-induced aggregation of DOC into what would become filterable POC¹⁹.

From 8 to 12 replicate mussels (mean shell length 19.5 ± 1.3 mm) were placed in 500 ml of DO¹⁴C solutions for 2 h at 16 °C. Exposure times were short to minimize recycling of ¹⁴C and aggregation of DOC. After exposures, mussel soft parts were dissected from shells, soft parts were dissolved in 0.75 ml Solvable (Packard), and counted by liquid scintillation²⁷. To correct for radioactivity derived from feeding water remaining in the mussels' mantle cavities, five additional mussels placed in each treatment were allowed to pump actively for 5 min and activities in these additional mussels were subtracted from activities in experimental mussels. Five min was considered sufficient time because each mussel of the size used here could filter 15–20 ml according to a published model¹⁷, sufficient to completely replace all water in the mantle with radioactive feeding water.

A separate batch of filtered Hudson River water was used to evaluate the influence of DOC on influx rates of radioisotopes of Ag, Cd, Se, and Hg in zebra mussels. Radioisotopes (^{110m}Ag, ¹⁰⁹Cd and ⁷⁵Se(IV) combined, and ²⁰³Hg independently) were added to water containing either 2.0 mg l⁻¹ LMW DOC, 2.9 mg l⁻¹ HMW DOC, or 3.3 mg l⁻¹ total DOC, allowed to equilibrate for 15 h, and presented to 7 replicate mussels in 500 ml for 2 h.

Radioisotope additions were 0.31–0.50 kBq ^{110m}Ag (in 0.1 M HNO₃), 0.79–0.99 kBq ¹⁰⁹Cd (in 0.1 M HCl), 22.5–29.3 kBq ⁷⁵Se (Na₂SeO₃ in distilled H₂O), and 21.9–34.1 kBq of ²⁰³Hg (in 1 M HCl); resulting concentrations were 0.38–0.49 nM for Ag, 0.21–0.22 nM for Cd, 0.94–1.0 nM for Se and 1.0–1.3 nM for Hg. Aliquots of water were periodically removed to monitor radioactivity and mussels were removed, dissected to remove shells, and gamma counted²⁷. Corrections were made for activity of water remaining in the mantle. Soft parts were dried (70 °C for 48 h) and weighed to estimate filtration rates from published equations¹⁷.

Received 27 March; accepted 6 June 2000.

- Strayer, D. L. *et al.* Arrival, spread, and early dynamics of a zebra mussel (*Dreissena polymorpha*) population in the Hudson River estuary. *Can. J. Fish. Aquat. Sci.* **53**, 1143–1149 (1996).
- Madon, S. P., Schneider, D. W., Stoeckel, J. A. & Sparks, R. E. Effects of inorganic sediment and food concentrations on energetic processes of the zebra mussel, *Dreissena polymorpha*: implications for growth in turbid rivers. *Can. J. Fish. Aquat. Sci.* **55**, 401–413 (1998).
- Fahnenstiel, G. L., Lang, G. A., Nalepa, T. F. & Johengen, T. H. Effects of zebra mussel (*Dreissena polymorpha*) colonization on water quality parameters in Saginaw Bay, Lake Huron. *J. Great Lakes Res.* **21**, 435–448 (1995).
- Caraco, N. F. *et al.* Zebra mussel invasion in a large, turbid river: phytoplankton response to increased grazing. *Ecology* **78**, 588–602 (1997).
- Findlay, S., Pace, M. L. & Fischer, D. T. Response of heterotrophic planktonic bacteria to the zebra mussel invasion of the tidal freshwater Hudson River. *Microb. Ecol.* **36**, 131–140 (1998).
- Stephens, G. C. Uptake of organic material by aquatic invertebrates. III. Uptake of glycine by brackish-water annelids. *Biol. Bull.* **126**, 150–162 (1964).
- Ferguson, J. C. A comparative study of the net metabolic benefits derived from the uptake and release of free amino acids by marine invertebrates. *Biol. Bull.* **162**, 1–17 (1982).
- Wright, S. H., Southwell, K. M. & Stephens, G. C. Autoradiographic analysis of amino acid uptake by the gill of *Mytilus*. *J. Comp. Physiol. B* **154**, 249–256 (1984).
- Wright, S. H. & Manahan, D. T. Integumental nutrient uptake by aquatic organisms. *Annu. Rev. Physiol.* **51**, 585–600 (1989).
- Wotton, R. S. Colloids, bubbles, and aggregates—a perspective on their role in suspension feeding. *J. N. Am. Benthol. Soc.* **15**, 127–135 (1996).
- Decho, A. W. Microbial exopolymer secretions in ocean environments: their role(s) in food webs and marine processes. *Oceanogr. Mar. Biol. Annu. Rev.* **28**, 73–153 (1990).
- Ciborowski, J. J. H., Craig, D. A. & Fry, K. M. Dissolved organic matter as food for black fly larvae (Diptera: Simuliidae). *J. N. Am. Benthol. Soc.* **16**, 771–780 (1997).
- Campbell, P. G. C. In *Metal Speciation and Bioavailability in Aquatic Systems* (eds Tessier, A. & Turner, D. R.) 45–102 (Wiley, Chichester, 1995).
- Köuch, J. & Pempkowiak, J. Molecular weight of humic acids as a major property of the substances influencing the accumulation rate of cadmium by a blue mussel (*Mytilus edulis*). *Environ. Int.* **22**, 585–589 (1996).
- Carvalho, R. A., Benfield, M. C. & Santschi, P. H. Comparative bioaccumulation studies of colloiddally complexed and free-ion heavy metals in juvenile brown shrimp *Penaeus aztecus* (Crustacea: Decapoda: Penaeidae). *Limnol. Oceanogr.* **44**, 403–414 (1999).
- Wang, W.-X., Fisher, N. S. & Luoma, S. N. Kinetic determinations of trace element bioaccumulation in the mussel *Mytilus edulis*. *Mar. Ecol. Prog. Ser.* **140**, 91–113 (1996).
- Kryger, J. & Riisgård, H. U. Filtration rate capacities in 6 species of European freshwater bivalves. *Oecologia* **77**, 34–38 (1988).
- Roditi, H. A. & Fisher, N. S. Rates and routes of trace element uptake in zebra mussels. *Limnol. Oceanogr.* **44**, 1730–1749 (1999).
- Kepkay, P. E. Particle aggregation and the biological reactivity of colloids. *Mar. Ecol. Prog. Ser.* **109**, 293–304 (1994).
- Cutter, G. A. Kinetic controls on the speciation of metalloids in seawater. *Mar. Chem.* **40**, 65–80 (1992).
- Howarth, R. W., Marino, R., Garritt, R. & Sherman, D. Ecosystem respiration and organic carbon processing in a large, tidally influenced river: the Hudson River. *Biogeochemistry* **16**, 83–102 (1992).
- Findlay, S., Pace, M. L. & Fischer, D. T. Spatial and temporal variability in the lower food web of the tidal freshwater Hudson River. *Estuaries* **19**, 866–873 (1996).
- Nichols, S. J. In *Zebra Mussels: Biology, Impacts, and Control* (eds Nalepa, T. F. & Schloesser, D. W.) 733–747 (Lewis, Boca Raton, 1993).
- Walz, N. The energy balance of the freshwater mussel *Dreissena polymorpha* Pallas in laboratory experiments and in Lake Constance. III. Growth under standard conditions. *Arch. Hydrobiol.* **55** (suppl.), 121–141 (1978).
- Fisher, N. S., Teyssie, J.-L., Fowler, S. W. & Wang, W.-X. Accumulation and retention of metals in mussels from food and water: a comparison under field and laboratory conditions. *Environ. Sci. Technol.* **30**, 3232–3242 (1996).
- Sañudo-Wilhelmy, S. A. & Gill, G. A. Phase-speciation of toxic metals in the Hudson River estuary. *Eos* **77**, F200 (1996).
- Lee, B.-G. & Fisher, N. S. Degradation and elemental release rates from phytoplankton debris and their geochemical implications. *Limnol. Oceanogr.* **37**, 1345–1360 (1992).
- Armstrong, F. A., Williams, P. M. & Strickland, J. D. Photooxidation of organic matter in seawater by ultraviolet radiation, analytical and other applications. *Nature* **211**, 481–487 (1966).
- Hedges, J. I. *et al.* Origins and processing of organic matter in the Amazon River as indicated by carbohydrates and amino acids. *Limnol. Oceanogr.* **39**, 743–761 (1994).
- Santschi, P. H., Guo, L., Means, J. C. & Ravichandran, M. In *Biogeochemistry of Gulf of Mexico Estuaries* (eds Bianchi, T. S., Pennock, R. & Twilley, R. R.) 347–380 (Wiley, New York, 1999).

Acknowledgements

We thank J. Cole, G. Stephens and D. Strayer for comments on this manuscript. This work was supported by the New York Sea Grant Institute, the National Association of Photographic Manufacturers, the Hudson River Foundation, and the National Science Foundation (N.F.), and a NOAA NERRS graduate fellowship (H.R.).

Correspondence and requests for materials should be sent to N.F. (e-mail: nfisher@notes.cc.sunysb.edu).

Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS

Shuji Shigenobu*†, Hidemi Watanabe†, Masahira Hattori†,
Yoshiyuki Sakaki†‡ & Hajime Ishikawa*

* Department of Biological Sciences, Graduate School of Science,
University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

† RIKEN Genomic Sciences Centre, Kitasato 1-15-1, Sagami-hara-shi, Kanagawa,
228-8555, Japan

‡ Human Genome Centre, Institute of Medical Science, University of Tokyo,
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Almost all aphid species (Homoptera, Insecta) have 60–80 huge cells called bacteriocytes, within which are round-shaped bacteria that are designated *Buchnera*¹. These bacteria are maternally transmitted to eggs and embryos through host generations, and the mutualism between the host and the bacteria is so obligate that neither can reproduce independently². *Buchnera* is a close relative of *Escherichia coli*³, but it contains more than 100 genomic copies per cell⁴, and its genome size is only a seventh of that of *E. coli*⁵. Here we report the complete genome sequence of *Buchnera* sp. strain APS, which is composed of one 640,681-base-pair chromosome and two small plasmids. There are genes for the biosyntheses of amino acids essential for the hosts in the genome, but those for non-essential amino acids are missing, indicating complementarity and syntrophy between the host and the symbiont. In addition, *Buchnera* lacks genes for the biosynthesis of cell-surface components, including lipopolysaccharides and phospholipids, regulator genes and genes involved in defence of the cell. These results indicate that *Buchnera* is completely symbiotic and viable only in its limited niche, the bacteriocyte.

One of the principal ecological niches of microbes is the inside of eukaryotic cells⁶. Although the majority of associations between microbe and eukaryote is either commensal or, quite often, mutually beneficial, most studies of animal-associated microbes have dealt with those rare bacterial species that cause diseases. Our major interest focuses on endocellular mutualistic bacteria, which, unlike pathogenic parasites, have been transmitted through host generations for an evolutionary length of time. The endocellular mutualistic associations must have evolved repeatedly and have had major consequences for the diversification of both bacteria and hosts. One typical mutualism is observed in the symbiosis between *Buchnera* and aphids. A phylogenetic analysis has indicated that the symbiotic relationship was established 200–250 Myr ago and led to co-speciation of the hosts and their symbionts⁷.

Buchnera sp. APS is harboured by the pea aphid, *Acyrtosiphon pisum* (Harris). We sequenced the *Buchnera* genome by the whole genome random sequencing method. The genome comprises one circular chromosome and two circular plasmids, pLeu and pTrp. The chromosome is 640,681 base pairs (bp), the smallest of the completely sequenced genomes, except for that of *Mycoplasma genitalium* (580,070 bp) which is regarded as a genome of the minimal gene set⁸. The pLeu plasmid is 7,786 bp and has 7 open reading frames (ORFs) including a *leuABCD* operon, and the pTrp plasmid has at least two tandem repeats of the *trpEG* operon^{3,9,10}. The average G+C content of the *Buchnera* genome is 26.3%. This AT richness is a feature of many endocellular bacteria, including both endosymbiotic and parasitic bacteria¹¹.

As *E. coli* and *Haemophilus influenzae* are closely related to *Buchnera* phylogenetically (see below), we assumed the DnaA box upstream of the *gidA* gene is the replication origin, and designated the start of the DnaA box as base pair one. *Buchnera* has only one

DnaA box, and slight shifts in the GC skew values are observed at the third codon position and in the non-coding region around *rho*, which is located 13 kilobases (kb) upstream of *gidA* (data not shown). We did not find an insertion sequence (IS) or a phage-related sequence by database homology search. Neither are there any significant repetitive elements. *Buchnera* has a single copy of each of the three types of ribosomal RNA and 32 transfer RNA genes.

We identified 583 ORFs in the genome with an average size of 988 bp, covering 88% of the whole genome (Fig. 1). It is intriguing that the predicted isoelectric points (pIs) of the products of the ORFs are on average much more basic than those of polypeptides of other bacteria. The average pI of *Buchnera* polypeptides is 9.6, whereas that of *E. coli* and *H. influenzae* polypeptides is 7.2 and 7.3,

Table 1 Gene repertoire of the *Buchnera* genome by functional category

Small-molecule metabolism	
Degradation of small molecules	3
Energy metabolism	51
Glycolysis	9
Pyruvate dehydrogenase	3
Tricarboxylic acid cycle	2
Pentose phosphate pathway	6
Entner–Doudoroff pathway	0
Respiration	23
Fermentation	0
ATP-proton motive force	8
Central intermediary metabolism	13
Amino-acid biosynthesis	55
Glutamate family	8
Aspartate family	11
Serine family	4
Aromatic amino acid family	16
Histidine	8
Pyruvate family	0
Branched-chain family	8
Polyamine biosynthesis	2
Purines, pyrimidines, nucleosides and nucleotides	34
Biosynthesis of cofactors, prosthetic groups and carriers	26
Fatty acid biosynthesis	6
Broad regulatory functions	
Broad regulatory functions	7
Macromolecule metabolism	
Synthesis and modification of macromolecules	187
rRNA and stable RNAs	4
Ribosomal protein synthesis and modification	54
Ribosome maturation and modification	0
tRNA	32
Aminoacyl tRNA synthetases and their modification	34
Nucleoproteins	2
DNA replication, restriction/modification and recombination	32
Protein translation and modification	18
RNA synthesis, RNA modification and DNA transcription	9
Polysaccharides (cytoplasmic)	0
Phospholipids	2
Degradation of macromolecules	22
Cell envelope	46
Membranes, lipoproteins and porins	4
Surface polysaccharides, lipopolysaccharides, and antigens	2
Surface structures	25
Murein sacculus and peptidoglycan	15
Cell processes	
Transport/binding proteins	18
Chaperones	11
Cell division	12
Chemotaxis and mobility	0
Protein and peptide secretion	7
Osmotic adaptation	0
Detoxification	3
Other	
Colicin-related functions	1
Drug/analogue sensitivity	2
Adaptations and atypical conditions	2
Other unclassified	111
Other categories	28
Conserved hypothetical	79
Unknown (unique to <i>Buchnera</i>)	4

All genes were classified according to Riley's classification³⁰. The number of genes for each category are listed. The complete list of genes is available on our website (<http://buchnera.gsc.riken.go.jp/>).

respectively. Comparison of amino-acid composition between *E. coli* and *Buchnera* shows that lysine usage of *Buchnera* is twice that of *E. coli*, causing the increased pI (data not shown). The 583 predicted ORFs were compared against a non-redundant protein database and their biological roles were assigned (Table 1). Similarity searching permitted the functional assignment of 500 ORFs, and another 79 ORFs are similar to hypothetical proteins deposited for other bacteria. Only four ORFs are unique to *Buchnera*. Generally, the most similar counterparts of *Buchnera* proteins are those of *E. coli*, and the gene order in *E. coli* operons is well conserved in *Buchnera*. Considering these observations, we conclude the *Buchnera* genome is a subset of the *E. coli* genome.

The biosynthesis capabilities of *Buchnera* characterize it as a symbiont. Endocellular and epicellular parasites that have dramatically reduced their genome size, like *Buchnera*, depend on their hosts for most nutrients, and the reduction of their genome size is, at least partly, due to the loss of biosynthetic genes for nutrients. However, nutritional and physiological studies show that *Buchnera* is a provider, rather than a recipient, of biosynthetic products

including essential amino acids and vitamins, to its host^{3,10,12–14}. We found 54 genes involved in amino-acid biosynthesis in the *Buchnera* genome. One of the most characteristic features of *Buchnera*, unveiled by our genome analysis, is that the genes for biosyntheses of the amino acids essential for the aphid hosts^{3,15} are present, but those for the non-essential amino acids are almost completely missing (Fig. 2). This complementarity of the gene repertoire shows how successfully the symbiosis is operating, in that *Buchnera* provides the host with what the host cannot synthesize, and conversely, the host provides the symbiont with what *Buchnera* cannot synthesize. Moreover, as the precursors of some essential amino acids are non-essential amino acids, glutamate and aspartate (Fig. 2a), the biosynthetic pathways of both the host and the symbiont are not only complementary, but also mutually dependent. This analysis is consistent with experimental evidence that aphids do not usually excrete a nitrogenous waste product, but recycle the amino groups as glutamine, which *Buchnera* uses as a substrate for the synthesis of essential amino acids^{10,13,16}.

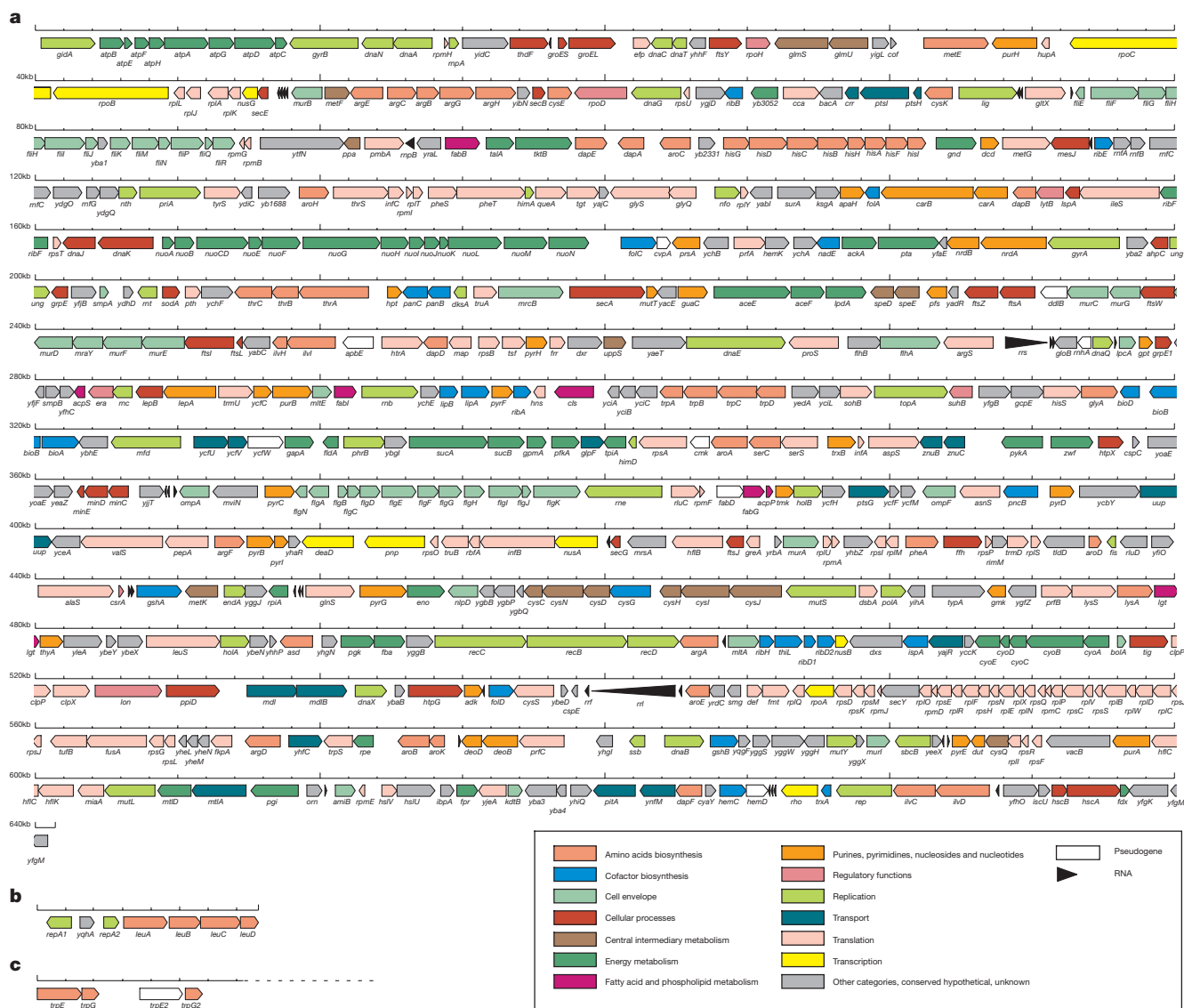


Figure 1 Linear representation of the *Buchnera* sp. APS genome illustrating the location of each predicted protein-coding region and RNA genes. All of these are essentially circular. **a**, Chromosome. **b**, pLeu plasmid. **c**, pTrp plasmid. The *trpEG* plasmid is composed of tandem repeats of the *trpEG* operon. As the repeats are highly similar, the

assembled sequence converged at 7,258 bp, which is equivalent to two units of the repeat. The pTrp plasmid of *Buchnera* of *A. pisum* may contain five, six or ten repeats of this operon³.

A similar example is observed in the pantothenate–coenzyme A (CoA) biosynthetic pathway. Although pantothenate seems to be synthesized from pyruvate in *Buchnera*, no genes for the pathway from pantothenate to CoA are found. On the other hand, animals generally have the ability to produce CoA from pantothenate, whereas they are not able to synthesize pantothenate itself. That *Buchnera* possesses complete gene sets for the sulphur reduction pathway and biosynthesis of cysteine is interesting, because insects cannot reduce sulphate to sulphide. Also experimental evidence shows that the *Buchnera*–bacteriocyte system is responsible for sulphate assimilation¹⁷. In contrast to obligatory parasitic bacteria, *Buchnera* has almost complete nucleotide biosynthetic pathways (Fig. 3). It is not known whether they are for the host or for its own use.

Such an obligatory mutualistic association as that between *Buchnera* and aphids should be maintained through exchange of various substances between the symbiont and the cytoplasm of the host cell. However, only a few transporter genes are present in the *Buchnera* genome (Table 1). Although the ABC transport system is a major class of cellular translocation machinery and many paralogous genes involved in this system are found in all bacterial species sequenced to date¹⁸, *Buchnera* has only a few ABC transporter genes.

Phosphoenolpyruvate–carbohydrate phosphotransferase systems (PTSs) seem to function in the *Buchnera* cell to import glucose and mannitol. Apart from these transporters, we found no other substrate-specific transporter genes. Hypothetical proteins YnfM and YajR are probably low-affinity transporters such as multidrug-efflux proteins. GlpF and OmpF-like porin may be involved in passive diffusion. The genes responsible for the *sec* protein secretion system are conserved in the *Buchnera* genome. Possibly, the flagellum of *Buchnera* serves as a transporter structure rather than a motor apparatus, as in *Salmonella typhimurium* and *Yersinia enterocolitica*^{19,20}. In general, the flagellum is composed of three components: a basal body, a hook and a filament. In the *Buchnera* genome, however, there is no evidence for a gene for filament (*fliC*), which confers motility on the cell. In addition, *Buchnera* lacks genes involved in chemotaxis. Indeed, neither flagellum nor motility has been observed with *Buchnera*.

The genome data indicates that *Buchnera* respire aerobically. This seems reasonable as this bacterium inhabits the bacteriocyte, which receives an ample supply of oxygen through the trachea and contains many mitochondria in the cytoplasm. *Buchnera* has complete gene sets responsible for glycolysis, the pentose phosphate cycle and aerobic respiration; however, it does not have a gene set for



Figure 2 Amino-acid biosynthetic pathways in *Buchnera* deduced from the gene set. The sequential pathways are represented by arrows, each of which indicates one step catalysed by the named enzyme. The steps for which no genes were found in the *Buchnera* genome are pink, as are precursors for which *de novo* synthetic pathways were not identified. **a**, Essential amino acids of animals. In the valine, isoleucine and leucine biosynthetic pathways, the gene for common enzyme *ilvE* is absent (purple). *ilvE* is a

branched-chain amino-acid aminotransferase and typically the final enzyme in the pathway, so the reaction may take place by some other aminotransferase enzyme. In the lysine biosynthetic pathway, the *dapC* gene (blue) has never been identified in any organisms. The terminal step of the phenylalanine biosynthesis is catalysed by TyrB in *E. coli*, but HisC may substitute for TyrB in *Buchnera*. **b**, Non-essential amino acids of animals. The alanine biosynthetic pathway has not been elucidated, even in *E. coli*.

operation of the TCA cycle apart from genes for the 2-oxoglutarate dehydrogenase complex. In the *Buchnera* genome, the NADH dehydrogenase (*nuo*) operon and the cytochrome *o* (*cyo*) operon are conserved with the same gene arrangements as *E. coli*, but the ubiquinone biosynthetic pathway is not even found. *Buchnera* has an F_0F_1 type ATP synthase operon, indicating that this bacterium is able to produce ATP using the proton electrochemical gradient generated by the electron transport system. *Buchnera* lacks genes responsible for fermentation and anaerobic respiration.

Buchnera seems to have a limited capacity for DNA repair and recombination, and exhibits an unusual repertoire of genes in this category. It is striking that the *recA* gene is missing from the *Buchnera* genome, as RecA is the most crucial component for the homologous recombination reaction. *Buchnera* is the first organism found to have *recBCD* without *recA*, though some mollicutes species have truncated *recA*²¹. Similarly, in the *uvr* excision repair system, *Buchnera* lacks *uvrABC*, but retains *mfd*. The *recA* and *uvrABC* are retained in all sequenced eubacterial genomes²², except for *Buchnera*. This unique inventory of repair genes implies that the repair system and the recombination mechanism of this symbiotic bacterium are severely impaired. Alternatively, *Buchnera* uses these

residual components differently from other organisms to provide the minimum requirement for its survival. The absence of a series of genes responsible for the SOS system, *recA*, *lexA*, *umuCD* and *uvrABC*, indicates that the *Buchnera* genome is vulnerable to DNA damage. Genes for DNA methylation and restriction are also missing, further evidence that *Buchnera* has limited defences.

Buchnera has only a few genes for cell-surface components. Our genome analysis indicates that *Buchnera* is not able to make lipopolysaccharides (LPSs). The genes for the biosynthesis of the LPS components, except for *lpcA* and *kdtB*, are missing from the *Buchnera* genome. We found only a few genes encoding lipoproteins and outer membrane proteins. Scarcity of genes for these components indicates that the cell surface of *Buchnera* is structurally vulnerable. This is in contrast to other bacteria, including pathogenic and free-living ones, which have complex and flexible surface structures to evade attack by the host immune system or to survive harsh environments. This structural fragility of *Buchnera* may be caused by its prolonged intracellular life, sheltered from attack by the host and foreign enemies.

Surprisingly, genes responsible for phospholipid biosynthesis are completely missing from the *Buchnera* genome, except that for

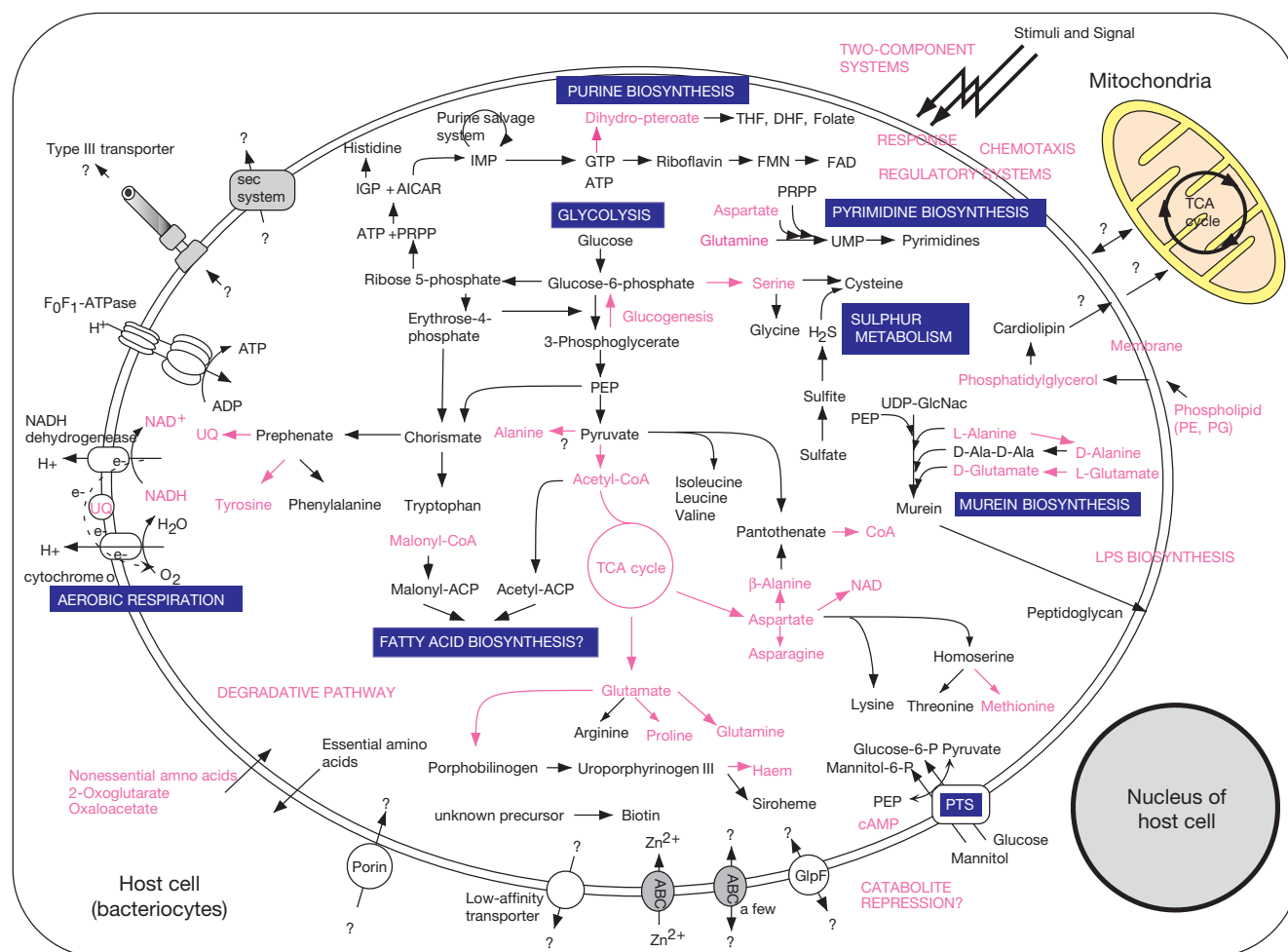


Figure 3 An integrated view of metabolism in *Buchnera* deduced from the genes identified. Pathways or steps for which no enzymes were identified are pink, as are the compounds for which *de novo* synthetic pathways were not identified. Question marks indicate that particular uncertainties exist or that the pathway has not been completely elucidated, even in *E. coli*. AICAR, 5'-phosphoribosyl-4-carboxamide-5-aminoimidazole;

cAMP, cyclic AMP; DHF, dihydrofolate; GlcNac, *N*-acetylglucosamine; IGP, imidazo-
leglycerol phosphate; IMP, inosine monophosphate; PE, phosphatidyl ethanolamine; PEP,
phosphoenolpyruvate; PG, phosphatidyl glycerol; PRPP, phosphoribosyl-pyrophosphate;
PTS, phosphoenolpyruvate:carbohydrate phosphotransferase system; TCA, tricarboxylic
acid; THF, tetrahydrofolate; UQ, ubiquinone.

cardiolipin synthetase (*cls*), although phospholipid is an indispensable component in the formation of the membrane lipid bilayer. Possibly, *Buchnera* either imports phospholipid from the host or synthesizes it, employing relevant enzymes transferred from the host cell, like mitochondria do.

Another prominent feature of the *Buchnera* genome is that genes for various regulatory systems are almost completely missing. Among these are two-component regulatory systems, which generally control gene expression in response to environmental changes. In addition, all the other types of transcriptional regulators, except *dnaA*, are missing. Indeed, no transcriptional regulator of amino-acid biosynthesis is present despite the conservation of many genes for amino-acid biosynthesis. Comparison of operon structure between *E. coli* and *Buchnera* indicates that genes of *Buchnera* do not have leader sequences, and that *Buchnera* is not equipped with a transcriptional attenuation system. Although *Buchnera* has PTSs, which are involved in catabolite repression through cyclic AMP in many bacteria, the genes for adenylate cyclase (*cyaA*) and cAMP receptor protein (*crp*) are absent, indicating the lack of transcriptional regulation for the response to carbon-source change. Instead, a carbon storage regulator CsrA might be involved in global post-transcriptional regulation of the carbohydrate metabolism²³ in *Buchnera*. The *Buchnera* genome contains only two predicted sigma factors, *rpoD* and *rpoH*. Other parasitic bacteria with small genomes, such as *M. genitalium* and *Rickettsia prowazekii*, have also lost large parts of their regulatory systems. However these organisms have also lost the genes that are normally under the control of the regulators. *Buchnera* is unique in lacking regulatory genes, but having their regulatees. It is possible that the loss of regulatory genes is a consequence of the homeostatic environment in which *Buchnera* has been housed for so long.

To evaluate the evolution of the characteristic gene set in *Buchnera*, we tried to reconstruct the history of *Buchnera*. First, we determined the evolutionary position of *Buchnera* among prokaryotes. We made orthologue groups of all ORFs in 23 complete prokaryotic genomes. For each group, we constructed a molecular phylogenetic tree and inferred the most plausible phylogeny of *Buchnera* and its relatives by searching for the most frequent sub-tree of the same topology that included one or more *Buchnera* genes or domains. The results indicate that after the speciation of *R. prowazekii*, *Buchnera* then diverged from the lineage to *E. coli* and *H. influenzae*, although a certain number of trees support the topology in which the closest relative of *E. coli* is *Buchnera*. To see whether the *Buchnera* genome is small because the genome of the last common ancestor (LCA) of *Buchnera*, *E. coli* and *H. influenzae* was as small as that of *Buchnera*, or because of gene loss in *Buchnera* after speciation, we inferred the gene set of the LCA. The gene set of *Buchnera*, excluding a few genes, is a small subset of that of the LCA, and many genes of the LCA were missing in *Buchnera*, such as those for non-essential amino-acid metabolism. In addition, no *Buchnera*-specific duplicated gene was found. These results strongly indicate that the small *Buchnera* genome is the result of reductive evolution.

The gene repertoire of the *Buchnera* genome is so specialized to intracellular life that it cannot survive outside the eukaryotic cell. Although the features of genome size reduction and heavy reliance on the host are shared between obligatory parasites and *Buchnera*, the gene sets show a marked difference in the manner of dependency between these two types of organisms. Whereas parasites depend upon nutrients from the host commensally, *Buchnera* provides nutrients to the host, using host-derived precursors. Moreover, *Buchnera* even seems to owe its membrane bilayer to the host and requires the host's protective environment. In this view, *Buchnera* is similar to organelles, which also require the endocellular environment of the host cells and make a contribution to the host, for example, in energy production.

This study is the first case where genomic evolution of a

mutualistic organism is revealed at the genomic level. Although this kind of organism has been difficult to study experimentally because of its absolute mutualism, this genomic data should promote experimental approaches to symbiosis, and further experimental data may give us an even deeper insight into the evolutionary significance of tight interspecies associations. □

Methods

Whole genome random shotgun sequencing

We prepared genomic DNA from the APS strain of *Buchnera* sp. harboured by the pea aphid, *Acyrtosiphon pisum* (Harris). *A. pisum* is a long-established parthenogenetic clone maintained on young broad bean plants, *Vicia faba* (L.), at 15 °C under a long-day regime of 18 h light and 6 h dark. We collected the bacteriocytes by dissecting 2,000 pea aphids in buffer A (35 mM Tris-HCl (pH 7.5), 25 mM KCl, 10 mM MgCl₂, 250 mM sucrose). The bacteriocytes were crushed by pipetting and subjected to filtration through a 5-µm filter to obtain *Buchnera*.

We prepared genomic DNA from *Buchnera* cells by a standard phenol/chloroform protocol. The shotgun sequence libraries were prepared as described²⁴, except that we used vectors with cohesive ends (one with a T-overhang and one partially filled) to avoid chimera formation. Briefly, the genomic DNA fragments were hydrodynamically sheared using HydroShear (GeneMachines), blunt-ended and subjected to A-tailing. The A-tailed fragments were ligated with pGEM-T Easy vector (Promega). The method of using partially filled-in restriction endonuclease fragments has been described²⁵. The sequences were assembled using PHRED (P. Green and B. Ewing, University of Washington) and PHRAP (P. Green, University of Washington) and the consensus sequence was checked and edited using CONSED (D. Gordon, University of Washington). The gaps between the contigs were closed by primer walking. About 7-fold genome coverage was achieved by 9,747 sequencing reactions. According to CONSED, the overall error probability was estimated at less than 0.01%. The sizes of the predicted restriction endonuclease fragments coincided with the physical map⁵.

Informatics

We used two strategies for identifying ORFs. An initial set of ORFs likely to encode proteins was identified by the GeneHacker program, a system for gene structure prediction using a hidden Markov model (HMM)²⁶. Both predicted ORFs and the intergenic regions were compared against a non-redundant protein database using the BLAST (D. Altschul, et al., NCBI) programs. Combining these results, we identified and annotated the ORFs. Frameshifts were detected and corrected where appropriate as described²⁷. The isoelectric point for each protein was predicted using the ISOELECTRIC program in the GCG analysis suite (Genetics Computer Group). Possible metabolic pathways were examined using the online service KEGG²⁸.

Comparative genomic analysis was performed, using 23 complete prokaryotic genomes: *Haemophilus influenzae* (NC_000907), *Mycoplasma genitalium* (NC_000908), *Methanococcus jannaschii* (NC_000909), *Synechocystis* sp. (NC_000911), *Mycoplasma pneumoniae* (NC_000912), *Helicobacter pylori* 26695 (NC_000915), *Helicobacter pylori* J99 (NC_000921), *Escherichia coli* (NC_000913), *Methanobacterium thermoautotrophicum* (NC_000916), *Bacillus subtilis* (NC_000964), *Archaeoglobus fulgidus* (NC_000917), *Borrelia burgdorferi* (NC_001318), *Aquifex aeolicus* (NC_000918), *Pyrococcus horikoshii* (NC_000961), *Pyrococcus abyssi* (NC_000868), *Mycobacterium tuberculosis* (NC_000962), *Treponema pallidum* (NC_000919), *Chlamydia trachomatis* (NC_000117), *Rickettsia prowazekii* (NC_000963), *Chlamydia pneumoniae* (NC_000922), *Aeropyrum pernix* (NC_000854), *Thermotoga maritima* (NC_000853) and *Buchnera* sp. APS. See Supplementary Information. These sequence data, except the *Buchnera* genome, were obtained from Entrez Genomes at NCBI website (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome>).

Paralogue and orthologue analyses were performed on the amino-acid sequences of all the predicted genes based on the BLAST bit scores. Species-specific duplicated genes were defined as paralogues of the same species that were more similar to each other than to any genes of other species. Orthologue groups were determined using an orthologue identification method²⁹, by taking into account gene duplications after speciation and multi-domain structures. For each orthologue group, the following analyses were carried out with our original script package, Whole genome Analysis Tools (WAT), for automated phylogenetic analyses: multiple alignment construction with CLUSTALW (J. Thompson et al., NCBI), extraction of conserved regions, using a program, xcons (H. W., unpublished), calculation of the distances between the group members with the PROTDIST program of the PHYLIP program package (J. Felsenstein, University of Washington), and construction of a neighbour-joining tree. Conserved regions were defined as the common segments in the CLUSTALW alignment, all possible pairs of which should show positive alignment scores using a local alignment scoring method.

The gene set in the LCA was inferred by a simple cladistical methodology. For a given tree topology, the gene set in the LCA of a clade is inferred from the assumption that the LCA had an ancestral gene for each orthologue group identified between the sister groups of the LCA or between descendant of the LCA and outgroup of the clade radiating from the LCA. This is a reasonable assumption as orthologous genes are defined as descendants of a single ancestral gene in the LCA inherited by different species, and an ancestral species should have had an ancestral orthologue corresponding to an orthologue pair identified between the descendant and the ancestor/outgroup. The actual gene set of the LCA may be larger than the one inferred as some genes could have been lost or evolved rapidly in either or both of the sister groups. The gene list of the LCA of *Buchnera*, *E. coli*, and *H. influenzae*, and the information on the WAT scripts used for this LCA analysis is available on request.

Received 11 January; accepted 31 May 2000.

- Munson, M. A., Baumann, P. & Kinsey, M. G. *Buchnera*, new genus and *Buchnera aphidicola*, new species, a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *Int. J. Syst. Bacteriol.* **41**, 566–568 (1991).
- Ishikawa, H. Biochemical and molecular aspects of endosymbiosis in insects. *Int. Rev. Cytol.* **116**, 1–45 (1989).
- Baumann, P., Moran, N. A. & Baumann, L. in *The Prokaryotes* (ed. Dworkin, M.) (Springer, New York, 2000).
- Komaki, K. & Ishikawa, H. Intracellular bacterial symbionts of aphids possess many genomic copies per bacterium. *J. Mol. Evol.* **48**, 717–722 (1999).
- Charles, H. & Ishikawa, H. Physical and genetic map of the genome of *Buchnera*, the primary endosymbiont of the pea aphid *Acyrtosiphon pisum*. *J. Mol. Evol.* **48**, 142–150 (1999).
- Buchner, P. *Endosymbiosis of Animals with Plant Microorganisms* (Wiley, New York, 1965).
- Moran, N. A., Munson, M. A., Baumann, P. & Ishikawa, H. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. R. Soc. Lond. B* **253**, 167–171 (1993).
- Fraser, C. M. *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403 (1995).
- van Ham, R. C., Moya, A. & Latorre, A. Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *J. Bacteriol.* **179**, 4768–4777 (1997).
- Douglas, A. E. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annu. Rev. Entomol.* **43**, 17–37 (1998).
- Moran, N. A. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl Acad. Sci. USA* **93**, 2873–2878 (1996).
- Houk, E. J. & Griffiths, G. W. Intracellular symbionts of the Homoptera. *Annu. Rev. Entomol.* **25**, 161–187 (1980).
- Sasaki, T. & Ishikawa, H. Production of essential amino acids from glutamate by mycetocyte symbionts of the pea aphid, *Acyrtosiphon pisum*. *J. Insect. Physiol.* **41**, 41–46 (1995).
- Nakabachi, A. & Ishikawa, H. Provision of riboflavin to the host aphid, *Acyrtosiphon pisum*, by endosymbiotic bacteria, *Buchnera*. *J. Insect. Physiol.* **45**, 1–6 (1999).
- Mittler, T. E. Dietary amino acid requirements of Aphid *Myzus persicae* affected by antibiotic uptake. *J. Nutr.* **101**, 1023–1028 (1971).
- Whitehead, L. F. & Douglas, A. E. A metabolic study of *Buchnera*, the intracellular bacterial symbionts of the pea aphid *Acyrtosiphon pisum*. *J. Gen. Microbiol.* **139**, 821–826 (1993).
- Douglas, A. E. Sulphate utilization in an aphid symbiosis. *Insect. Biochem. Mol. Biol.* **18**, 599–605 (1988).
- Tomii, K. & Kanehisa, M. A comparative analysis of ABC transporters in complete microbial genomes. *Genome Res.* **8**, 1048–1059 (1998).
- Kubori, T. *et al.* Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science* **280**, 602–605 (1998).
- Young, G. M., Schmiel, D. H. & Miller, V. L. A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc. Natl Acad. Sci. USA* **96**, 6456–6461 (1999).
- Marais, A., Bove, J. M. & Renaudin, J. Characterization of the *recA* gene regions of *Spiroplasma citri* and *Spiroplasma melliferum*. *J. Bacteriol.* **178**, 7003–7009 (1996).
- Eisen, J. A. & Hanawalt, P. C. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.* **435**, 171–213 (1999).
- Nogueira, T. & Springer, M. Post-transcriptional control by global regulators of gene expression in bacteria. *Curr. Opin. Microbiol.* **3**, 154–158 (2000).
- Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
- Hattori, M. *et al.* A novel method for making nested deletions and its application for sequencing of a 300 kb region of human APP locus. *Nucleic Acids Res.* **25**, 1802–1808 (1997).
- Yada, T. & Hirose, M. Detection of short protein coding regions within the cyanobacterium genome: application of the hidden Markov model. *DNA Res.* **3**, 355–361 (1996).
- Tomb, J. F. *et al.* The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547 (1997).
- Ogata, H. *et al.* KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
- Watanabe, H., Mori, H., Itoh, T. & Gojobori, T. Genome plasticity as a paradigm of eubacteria evolution. *J. Mol. Evol.* **44** (Suppl. 1), S57–S64 (1997).
- Riley, M. Functions of the gene products of *Escherichia coli*. *Microbiol. Rev.* **57**, 862–952 (1993).

Supplementary information is available on *Nature's* World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

We thank the technical staff of RIKEN GSC, M. Horishima, H. Ishizaki, N. Ota and Y. Seki for sequencing; T. Yada for ORF prediction; A. Toyoda for technical support on the sequence library preparation; C. Kawagoe for computer system support; and T. D. Taylor for discussion. This work was supported by a grant from the Program for Promotion of Basic Research Activities for Innovation Biosciences (ProBRAIN) of the Bio-oriented Technology Research Advancement Institution, and Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Science, Sports and Culture.

Correspondence and requests for materials should be addressed to Y.S. (e-mail: sakaki@ims.u-tokyo.ac.jp) or H.I. (e-mail: iskw@biol.s.u-tokyo.ac.jp). The complete sequence and the annotated data are available on our website (<http://buchnera.gsc.riken.go.jp/>). The sequence has been deposited with DDBJ under accession number AP000398, AP001070 and AP001071 for chromosome, the pTnp plasmid and the pLeu plasmid, respectively.

Cloned pigs produced by nuclear transfer from adult somatic cells

Irina A. Polejaeva*, Shu-Hung Chen*, Todd D. Vaught*, Raymond L. Page*, June Mullins*, Suyapa Ball*, Yifan Dai*, Jeremy Boone*, Shawn Walker*, David L. Ayares*, Alan Colman† & Keith H. S. Campbell†‡

* PPL Therapeutics Incorporated, 1700 Kraft Drive, Blacksburg, Virginia 24060, USA

† PPL Therapeutics, Roslin, Midlothian EH25 9PP, UK

Since the first report of live mammals produced by nuclear transfer from a cultured differentiated cell population in 1995 (ref. 1), successful development has been obtained in sheep^{2,3}, cattle⁴, mice⁵ and goats⁶ using a variety of somatic cell types as nuclear donors. The methodology used for embryo reconstruction in each of these species is essentially similar: diploid donor nuclei have been transplanted into enucleated MII oocytes that are activated on, or after transfer. In sheep² and goat⁶ pre-activated oocytes have also proved successful as cytoplasm recipients. The reconstructed embryos are then cultured and selected embryos transferred to surrogate recipients for development to term. In pigs, nuclear transfer has been significantly less successful; a single piglet was reported after transfer of a blastomere nucleus from a four-cell embryo to an enucleated oocyte⁷; however, no live offspring were obtained in studies using somatic cells such as diploid or mitotic fetal fibroblasts as nuclear donors^{8,9}. The development of embryos reconstructed by nuclear transfer is dependent upon a range of factors. Here we investigate some of these factors and report the successful production of cloned piglets from a cultured adult somatic cell population using a new nuclear transfer procedure.

To date, the efficiency of somatic cell nuclear transfer, when measured as development to term as a proportion of oocytes used, has been very low (1–2%)¹⁰. A variety of factors probably contribute to this inefficiency. These include laboratory to laboratory variation, oocyte source and quality, methods of embryo culture (which are more advanced in some species (such as cows) than others (such as pigs)), donor cell type, possible loss of somatic imprinting in the nuclei of the reconstructed embryo, failure to reprogram the transplanted nucleus adequately, and finally, the failure of artificial methods of activation to emulate reproducibly those crucial membrane-mediated events that accompany fertilization.

In the pig, there is the additional difficulty that several (> 4) good quality embryos are required to induce and maintain a pregnancy¹¹. As fully developmentally competent embryos are rare in nuclear transfer procedures, there is every chance of squandering those good embryos unless very large numbers of reconstructed embryos are transferred back into recipients. Even if it were possible in the pig to select good quality blastocysts for transfer (after, for example, the use of a temporary recipient), most blastocysts formed from reconstructed embryos in other species are not competent to proceed to term¹⁰. The co-transfer of reconstructed embryos with 'helper', unmanipulated embryos, parthenotes or tetraploid embryos has been suggested as an aid to inducing and maintaining pregnancy. However, studies in mice after zygote pronuclear injection have suggested that the manipulated embryos are 'compromised' and selected against¹². An alternative to the use of 'helper' embryos is the hormonal treatment of recipient sows to maintain pregnancy with low embryo numbers¹³.

We cannot currently address all of the methodological problems, and, to improve our chances of success in pig nuclear

‡ Present address: University of Nottingham, School of Biosciences, Sutton Bonington, Leicestershire LE12 5RN, UK.

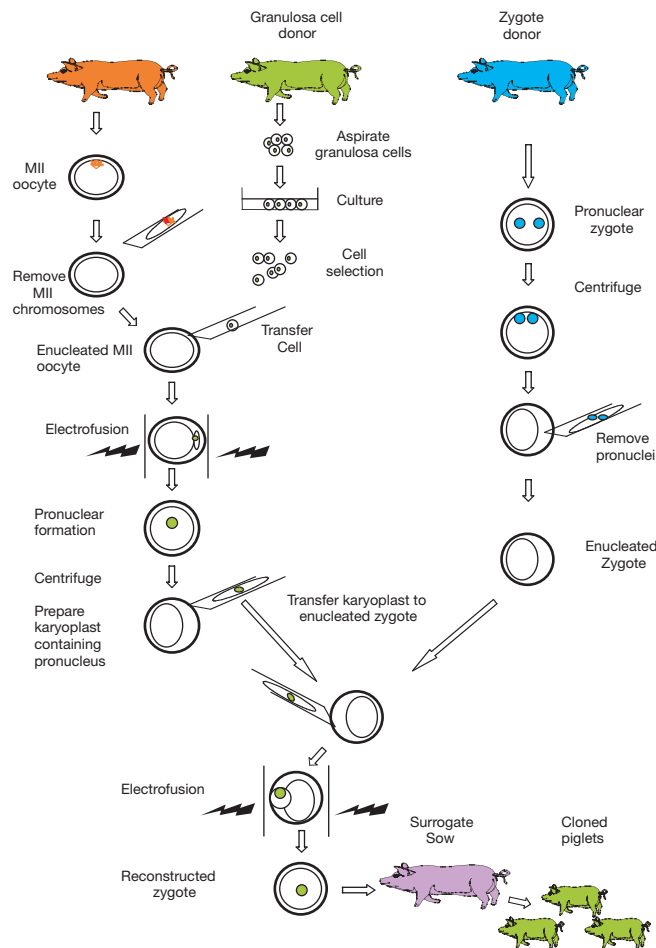


Figure 1 Representation of the double nuclear transfer procedure for the production of viable piglets using cultured adult somatic granulosa cells as nuclear donors. The outer circle in all the oocytes and embryos denotes the zona pellucida; the inner circle denotes the cell membrane.

transfer, we chose to focus on four areas: activation, choice of donor cell, embryo culture, and induction and maintenance of pregnancy.

In all species, when using MII oocytes as recipients, the method of activation is crucial for subsequent development. In the pig, although current activation protocols stimulate pronuclear



Figure 2 A litter of five live piglets derived by nuclear transfer using cultured adult granulosa cells as nuclear donors. A total of 72 reconstructed embryos were transferred to the surrogate sow.

formation, cleavage, and development to the blastocyst stage, both the frequency of development and the quality of the embryos produced are low¹⁴. A system that involves the use of fertilized zygotes as cytoplasmic recipients would bypass the inefficiencies of artificial activation procedures and might promote more successful development. The technique of pronuclear exchange between zygotes showed that the manipulations involved were compatible with development¹⁵; however, when donor nuclei from later developmental stages were transferred there was restricted development¹⁶. One explanation is that factors required for development, which are absent in the donor nuclei, are removed with the pronuclei. But if a pronucleus-like structure could be produced from the donor nucleus, this might prove a suitable nuclear donor for transfer. Such a system was described in mice by Kwon and Kono¹⁷, who first fused mitotically arrested blastomere nuclei with enucleated MII oocytes. The reconstructed oocytes were subsequently activated in the presence of cytochalasin B, preventing polar body extrusion and resulting in the formation of two diploid pseudo-pronuclei. Each pseudo-pronucleus was then transferred into an enucleated, *in vivo* produced zygote, which was transferred into a surrogate recipient for development to term. Effectively, this latter procedure mimics pronuclear exchange and allows the

Table 1 Development of porcine embryos

Cell isolate	Double NT							Single NT		
	Pool1	GR5	GR8	GR12	GR21Z	GR18	GR18	GR1	GR8	GR18
Cell treatment	CI	CI	CI	CI	CI	SS	SS	CI	CI	SS
No. of oocytes	245	344	269	291	311	(74)	(51)	123	109	(51)
No. of attempted reconstructions day 1 (%)	183 (75)	217 (63)	207 (77)	226 (78)	221 (71)	122 (63)	192 (89)	94 (76)	83 (76)	N/A
No. of fused embryos day 1 (%)	124 (68)	153 (70)	186 (90)	97 (43)	163 (74)	90 (74)	162 (84)	87 (93)	61 (73)	N/A
No. of day-1 embryos with single pronucleus (%)	87 (70)	88 (57)	120 (64)	62 (64)	69 (42)	23 (26)	102 (63)	—	—	—
No. of reconstructed embryos day 2	74	57	105	61	55	22	45*	—	—	—
No. of fused embryos day 2 (%)	72 (97)	56 (98)	100 (95)	53 (87)	54 (98)	22 (100)	44 (98)	—	—	—
No. of embryos transferred to recipient	72	56	100	53	54	22	44	85	61	39†
Pregnancy (no. of fetuses observed)	+ve (3)	—ve	—ve	—ve	—ve	+ve (6)	—ve	—ve	—ve	—ve
No. of live births (%)	5 (7)	0	0	0	0	0	0	0	0	0

Development of porcine embryos reconstructed using a single or double nuclear transfer protocol with adult granulosa cells cultured to confluence (CI), or serum starved (SS) as nuclear donors.

* Insufficient zygotes to reconstruct day 1 embryos.

† Due to insufficient zygote numbers 39 day-1 reconstructed embryos were transferred to a single recipient.

formation of a final reconstructed one-cell embryo whose membrane has been activated during fertilization.

The use of cultured cell populations for the production of animals by nuclear transfer is now well documented in a number of species. We have considerable experience in the production of sheep and cattle from primary cell populations and genetically modified primary cell populations. Analysis of these studies has shown considerable variation in development between individual cell populations and at present has provided no definitive method for the identification of cell populations that are suitable for nuclear transfer. Factors that are thought to influence the suitability of a particular cell population include the effects of oxidative damage associated with cellular metabolism, genome instabilities and chromosomal pathologies. All of these factors may be influenced by the method of isolation and culture, and the number of population doublings in culture. On consideration of these factors and our previous observations, we chose to use granulosa cells as nuclear donors. Granulosa cells are suitable nuclear donors in cattle¹⁸, and require the minimum of manipulations to establish in culture. Because of differences between cell populations, we initially decided to use a pool of cells isolated from a group of four donors. In later experiments, cell populations from individual animals were also examined. To minimize the culture period, early passage, never-frozen cells were used.

For embryo reconstruction, we attempted to minimize the potential inefficiencies at each step of the nuclear transfer procedure and adopted an approach that (1) uses *in vivo* derived material, (2) seeks to avoid artificial activation, and (3) minimizes the period of *in vitro* culture of manipulated embryos. To do this we used a two-stage nuclear transfer procedure modified from Kwon and Kono¹⁷ (Fig. 1). In the first stage, donor cells were fused to *in vivo* derived, enucleated MII oocytes obtained from superovulated crossbred gilts. The pseudo-pronucleus formed in the first nuclear transfer embryo was then subsequently transplanted into an *in vivo* produced, enucleated zygote (second nuclear transfer embryo). The second nuclear transfer reconstructed embryo was transferred to the oviduct of a synchronized sow within 2 h of fusion. Because of the expected low developmental rate, we transferred up to 100 reconstructed embryos to a single recipient. Each recipient was treated with pregnant mare serum gonadotropin (PMSG) and human

chorionic gonadotropin (hCG) to maintain pregnancy¹³ in the event that fewer than four reconstructed embryos were viable at implantation.

Coordination of the cell-cycle stages of the recipient cytoplasm and the donor nucleus are essential for maintaining correct ploidy and preventing DNA damage in nuclear transfer reconstructed embryos¹⁹. Various combinations of donor and recipient cell-cycle stages can prevent DNA damage and uncoordinated DNA replication, and result in formation of a pseudo-pronucleus. It has been suggested that the use of MII oocytes may improve 're-programming' of the donor genetic material owing to the occurrence of nuclear envelope breakdown and premature chromosome condensation, thus exposing the donor chromatin to maternally derived oocyte factors involved in early development. To take advantage of this here, we used MII oocytes as cytoplasm recipients for the first nuclear transfer embryo reconstruction. To maintain ploidy in this situation, we chose diploid donor nuclei as nuclear donors.

Previous studies have suggested that diploid cells arrested in the G0 phase of the cell cycle may be beneficial². Using flow cytometry, we examined the cell-cycle distribution of porcine granulosa cells under three different culture conditions: sub-confluent actively growing, 100% confluent, and cells starved of serum for 48 hours (see Figure in Supplementary Information). After serum starvation, the population contained a large proportion (7.2%) of cells with a DNA content lower than that consistent with a diploid cell (termed sub-G1). In contrast, in the population synchronized by contact inhibition, 90.3% of the cells had a diploid DNA content (G1/G0) and there were fewer sub-G1 cells (1.6%). We analysed DNA synthesis in serum-starved and contact-inhibited cell populations by 5-bromo-2'-deoxyuridine (BrdU) incorporation. These experiments revealed that 45% of the contact inhibited cell population compared with 0% of the serum-starved population incorporated BrdU. An analysis of BrdU incorporation after an additional 24 h of contact inhibition revealed that the fraction of BrdU-positive cells was reduced to 5%. These observations suggest that the diploid cells in the contact-inhibited granulosa cell population used as nuclear donors for embryo reconstruction contained a mixture of cell-cycle-arrested diploid cells (G1/G0) and unarrested diploid cells (G1), which were able to undergo a further round of DNA synthesis. In

Table 2 Microsatellite analysis of pigs and cell donors

Loci	S0059	S0070	S0122	S0226	SW24	SW72	SW840	SW936	TNFB
Samples									
PGR1	152	275	178	178	103	102	129	95	161
	152	295	182	198	111	110	129	97	164
PGR2	148	275	180	178	103	110	129	97	158
	154	275	182	198	111	112	129	111	185
PGR3	146	275	180	178	103	110	125	97	N/A
	152	275	182	198	109	112	125	111	
PGR4	152	275	178	198	103	102	129	97	158
	152	295	182	198	111	110	129	109	161
NTP1	152	275	178	178	103	102	129	95	161
	152	295	182	198	111	110	129	97	164
NTP2	152	275	178	178	103	102	129	95	161
	152	295	182	198	111	110	129	97	164
NTP3	152	275	178	178	103	102	129	95	161
	152	295	182	198	111	110	129	97	164
NTP4	152	275	178	198	103	102	129	97	158
	152	295	182	198	111	110	129	109	161
NTP5	152	275	178	198	103	102	129	97	158
	152	295	182	198	111	110	129	109	161
Recipient (54B)	152	275	178	192	111	102	N/A	103	158
	152	295	182	198	95	102		109	161
Boar	134	265	178	180	115	102	129	97	164
	156	273	182	180	115	112	129	109	185

Microsatellite analysis was performed on genomic DNA from the four individual populations of granulosa cells (PGR1, PGR2, PGR3, PGR4), the piglets (NTP1–5), the surrogate sow (54B) and the boar responsible for inseminating the zygote donors. Primers corresponding to nine polymorphic loci were used. Two numbers are shown for each sample at each locus, which represent the PCR product size for each of the two alleles at that particular locus.

contrast, when the serum-starved populations were used as nuclear donors most the diploid cells were cell-cycle arrested (G1/G0).

Production of the first nuclear transfer embryos requires activation of the *in vivo* derived oocytes. Activation experiments carried out in control oocytes showed that electrical stimulation applied between 51.5 and 60 h after hCG administration, promoted similar cleavage and development to blastocyst (see Table in Supplementary Information). For embryo reconstruction, MII oocytes were collected 46–54 h after hCG and the first nuclear transfer embryo reconstruction was carried out between 50 and 58 h after hCG. Reconstructed embryos were cultured overnight in NCSU-23 medium²²; we then checked them for the presence of a pronucleus and used them for the second nuclear transfer embryo reconstruction. The development of single nuclear transfer and double nuclear transfer embryos reconstructed from contact-inhibited and serum-starved granulosa populations were compared (see Table 1). In total, 185 single nuclear transfer embryos were transferred to 3 recipient sows and 401 double nuclear transfer embryos to 7 recipients. Two recipients of the double nuclear transfer embryos became pregnant as determined by ultrasound visualization of fetuses at day 35 of gestation. One of these maintained the pregnancy to term, and five piglets (Fig. 2) were delivered by Caesarean section on day 116 of gestation. The average birth weight of the piglets was 2.72 lb (range 2.28–3.08 lb); this is about 25% lower than that observed in the same population of pigs under natural mating conditions (average litter size average 10.9, average birth weight 3.6 lb, range 3.3–3.9 lb).

The live piglets were produced from a pooled population of cells derived from four animals. We carried out microsatellite analysis of genomic DNA from the various samples (Table 2). The comparison of the pattern of alleles in the piglets with that of the granulosa cell populations indicated that three of the nuclear transfer piglets (NTP1, NTP2 and NTP3) were derived from the porcine granulosa (PGR)1/cell line, as there was 100% identity at all nine microsatellite markers. The other 2 nuclear transfer piglets (NTP4 and NTP5) showed perfect identity with the genotype of the PGR4 cell population. All five of the nuclear transfer piglets were significantly different from the surrogate mother (54B). Some of the loci (S0059, S0070, S0122 and TNFB) were not highly polymorphic indicating a degree of homogeneity or inbreeding within the population of pigs used in these studies (all of which come from the same commercial supplier).

We think that the principal reasons for the success of this modified nuclear transfer procedure in pigs is its lack of reliance on current artificial activation protocols and *in vitro* culture techniques. Although elaborate, the double nuclear transfer does not add another major inefficiency (the second step fusion is very efficient). Direct transfer of a somatic nucleus to an enucleated zygote will not work because (in addition to reprogramming difficulties) of the loss of important factors sequestered within the removed pronuclei. The cell population used successfully as nuclear donors in these experiments were not quiesced by serum starvation. Cell-cycle analysis showed that most cells in control cultures had a diploid DNA content, and a high percentage were able to undergo a further round of DNA synthesis suggesting that most cells in the population were in the G1 phase of the cell cycle and not arrested in G0. All five of the pigs, now three months old, are extremely healthy, in contrast to the (usual) 50% postnatal loss of nuclear transfer animals¹⁰. It is tempting then to speculate that this modified method may have general utility in other species, even those where single nuclear transfer has been shown to work.

The successful development of nuclear transfer in pigs opens the door for the application of gene-targeting technology, thus allowing for very precise genetic modifications, including gene knockouts. We have recently reported gene targeting in cultured ovine somatic cells and the successful development to term of offspring produced by nuclear transfer using these cells²⁰. In pigs, a gene of great interest for the application of knockout technology

is that for α -1,3-galactosyl transferase (α -1,3-GT)—the enzyme responsible for adding the xenogeneic sugar, galactose α -1,3-galactose, to the surface of porcine cells. This gene is inactive in certain monkeys and humans, and their blood contains anti-gal antibodies, which trigger (in monkeys) early rejection of transplanted organs²¹. We have achieved targeted disruption of the α -1,3-GT gene in primary porcine cells (unpublished data) and this will allow the production of α -1,3-GT-deficient pigs, whose organs should show improved resistance to rejection. Overcoming antibody-mediated rejection is the first critical step in improving xenograft survival, towards the ultimate goal of providing an unlimited supply of compatible pig organs for human transplantation. □

Methods

Modified NCSU-23 medium

The published NCSU-23 medium²² was modified for use as a phosphate-buffered benchtop medium without NaHCO₃. Physiological pH phosphate buffer is made using a 3:1 molar ratio of dibasic to monobasic phosphate anions. These changes induced alterations in the Na and K concentrations, which were corrected by adjusting the NaCl and KCl concentrations to maintain osmolality and Na/K ratio (all chemicals purchased from Sigma unless otherwise noted).

Superovulation of donor gilts for collection of oocytes and zygotes

Crossbred gilts (280–320 lbs) were synchronized by oral administration of 18–20 mg Regu-Mate (Altrenogest, Hoechst) mixed into the feed. Regu-Mate was fed for 5–14 d using a scheme dependent on the stage of the oestrous cycle. Estrumate (250 µg, Bayer) was administered intramuscularly (i.m.) on the last day of the Regu-Mate treatment. Superovulation was induced with a single i.m. injection of 1,500 IU of PMSG (Diosynth) 15–17 h after the last Regu-Mate feeding. One thousand units of hCG (Intervet America) were administered i.m. 82 h after the PMSG injection.

We collected oocytes 46–54 h after the hCG injection by reverse flush of the oviducts using pre-warmed Dulbecco's phosphate buffered saline (PBS) containing bovine serum albumin (BSA; 4 g l⁻¹). For the collection of zygotes, 24–36 h after the hCG injection the gilts were either artificially inseminated or bred naturally. We flushed zygotes from the oviduct 52–54 h after the hCG injection using PBS containing BSA (4 g l⁻¹).

Isolation and culture of porcine granulosa cells

Follicular fluid was aspirated from 2–8-mm diameter follicles of superovulated crossbred gilts (Large White (1/2), Landrace (1/4), White Duroc (1/4)), 7–8 months old, 280–320 lb, 28–51 h post hCG injection. Granulosa cells were collected by centrifugation at 1,040 g for 10 min, re-suspended in DMEM (Gibco), containing 10% fetal calf serum (FCS; Summit Biotech), 0.1 mM non-essential amino acids (NEAA Gibco), 2 ng ml⁻¹ basic fibroblast growth factor (bFGF) (Beckton Dickinson) and 6 µl ml⁻¹ Gentamycin (Sigma). Cells were expanded for several days and then cryo-preserved.

For nuclear transfer, we plated the granulosa cells at 1–5 × 10⁴ cells per 35 mm dish in DMEM medium supplemented with NEAA (0.1 mM), bFGF (2 ng ml⁻¹) and 10% FCS, and cultured them to 100% confluency at 37 °C. For experiments where serum starvation was evaluated, cells were starved of serum for 48–72 h in DMEM containing 0.5% FCS. We collected cells by trypsinization and stored them in suspension in modified NCSU-23 phosphate medium at 38.5 °C for 20–120 min, before use as nuclear donors.

Activation of oocytes

Activation of control oocytes was achieved by application of two 1.0 kV cm⁻¹ DC electric pulses for 60 µs each at an interval of 5 s in activation medium (0.3 M D-sorbitol supplemented with 0.1 mM Mg SO₄ and 0.05 mM CaCl₂ in H₂O).

Reconstruction of first nuclear transfer embryo

Recovered oocytes were washed in PBS containing 4 g l⁻¹ BSA at 38 °C, and transferred to calcium-free phosphate-buffered NCSU-23 medium at 38 °C for transport to the laboratory. For enucleation, we incubated the oocytes in calcium-free phosphate-buffered NCSU-23 medium containing 5 µg ml⁻¹ cytochalasin B (Sigma) and 7.5 µg ml⁻¹ Hoechst 33342 (Sigma) at 38 °C for 20 min. A small amount of cytoplasm from directly beneath the first polar body was then aspirated using an 18-µm glass pipette (Humagen, Charlottesville, Virginia). We exposed the aspirated karyoplast to ultraviolet light to confirm the presence of a metaphase plate. A single granulosa cell was placed below the zona pellucida in contact with each enucleated oocyte. The couplet was transferred to a fusion chamber (model no. BT-453, BTX Inc., San Diego) containing 700 µl of 0.3 M mannitol, 0.1 mM MgSO₄ and 0.1 mM CaCl₂ in deionized water. Fusion and activation were induced by application of an AC pulse of 5 V for 5 s followed by two DC pulses of 1.5 kV cm⁻¹ for 60 µs using an ECM2001 Electroculture Manipulator (BTX Inc., San Diego). Couplets were then washed in bicarbonate buffered NCSU-23 medium, and incubated in this medium for 0.5–1 h at 38.6 °C in a humidified atmosphere consisting of 5% CO₂ in air. We checked couplets for fusion at ×300 magnification using an inverted microscope. Fused embryos were given a second activation stimulus of two successive DC pulses of 1.2 kV cm⁻¹ for 60 µs each, and cultured overnight in NCSU medium²⁰.

Reconstruction of second nuclear transfer embryo

Zygotes were centrifuged at 14,900 g for 15 min in a Biofuge 13 centrifuge and then incubated in phosphate buffered NCSU-23 medium containing $5.0 \mu\text{g ml}^{-1}$ cytochalasin B (Sigma) at 38°C for 20 min. Zygotes containing two pronuclei were enucleated using a 25–35- μm glass pipette by aspirating a membrane bound karyoplast containing both pronuclei (and the second polar body if present). Karyoplasts containing the pseudo-pronucleus were prepared from the day 1 nuclear transfer embryos as described for zygote enucleation with the modification that a 30–45- μm enucleation pipette was used for manipulation. A single karyoplast was placed into the perivitelline space of each enucleated zygote. Fusion was induced by application of an AC pulse of 5 V for 5 s followed by two DC pulses of 1.2 kV cm^{-1} for 60 μs . Couplets were then washed and cultured in NCSU-23 medium for 0.5–1 h at 38.6°C in a humidified atmosphere of 5% CO_2 . We transferred fused couplets as soon as possible to the oviduct of an oestrus-synchronized recipient gilt.

Treatment of recipient sows

Pregnancy was maintained by using a combination of PMSG and hCG. PMSG (1,000 IU) was injected i.m. on day 10 of the oestrous cycle (day 1 being the day of oestrus), hCG was injected i.m. 3–3.5 d later (day 13 of the cycle)¹⁵.

Microsatellite analysis

DNA ($25 \text{ ng } \mu\text{L}^{-1}$) from each of the five piglets (24-h tail samples), the four granulosa cell lines mixed to make pool 1 (PGR1, PGR2, PGR3 and PGR4), the recipient sow (54B) and a boar, 'Ranger', that was used for artificial insemination purposes, were sent in individually coded vials to Celera-AgGEN (Davis, California), a company that specializes in parentage verification for swine. The microsatellite analysis consists of a multiplexed set of nine polymorphic porcine loci, each of which consists of different multimers of short tandem repeats (dinucleotides). The polymorphic loci are designated: S0059, S0070, S0122, S0226, SW24, SW72, SW840, SW936 and TNFB (PCR primer sequence information is proprietary to Celera-AgGEN). This multiplex set contains nine different PCR primer pairs, amplified in two PCR reactions, a five-plex and a four-plex. Ten nanograms of template DNA was amplified in each multiplex PCR. The forward primers were labelled on the 5' end with a fluorescent dye (either FAM, JOE or TAMRA). The entire battery for each DNA sample was loaded in a single lane and co-electrophoresed with the internal size standard GeneScan 350 Rox. All multiplexing and loci evaluations were performed on an ABI PRISM 377 DNA Sequencer and analysed with Genotyper 2.0 software.

Received 12 June; accepted 27 July 2000.

- Campbell, K. H., McWhir, J., Ritchie, W. A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380**, 64–67 (1996).
- Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813 (1997).
- Wells, D. N., Misica, P. M., Day, T. A. & Tervit, H. R. Production of cloned lambs from an established embryonic cell line: a comparison between in vivo- and in vitro-matured cytoplasts. *Biol. Reprod.* **57**, 385–393 (1997).
- Cibelli, J. B. *et al.* Cloned transgenic calves produced from nonquiescent fetal fibroblasts. *Science* **280**, 1256–1258 (1998).
- Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R. & Yanagimachi, R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394**, 369–374 (1998).
- Baguissi, A. *et al.* Production of goats by somatic cell nuclear transfer. *Nature Biotechnol.* **17**, 456–461 (1999).
- Prather, R. S., Sims, M. M. & First, N. L. Nuclear transplantation in early pig embryos. *Biol. Reprod.* **41**, 414–418 (1989).
- Tao, T. *et al.* Development of pig embryos reconstructed by microinjection of cultured fetal fibroblast cells into in vitro matured oocytes. *Anim. Reprod. Sci.* **56**, 133–141 (1999).
- Ouhibi, N. *et al.* Nuclear transplantation in pigs: M-phase karyoplast to M-phase cytoplast fusion. *Reprod. Nutr. Dev.* **36**, 661–666 (1996).
- Colman, A. Somatic cell nuclear transfer in mammals: Progress and applications. *Cloning* **1**, 185–200 (2000).
- Polge, C., Rowson, L. E. A. & Chang, M. C. The effect of reducing the number of embryos during early stages of gestation on the maintenance of pregnancy in the pig. *J. Reprod. Fert.* **12**, 395–397 (1966).
- Canseco, R. S. *et al.* Gene transfer efficiency during gestation and the influence of co-transfer of non-manipulated embryos on production of transgenic mice. *Transgenic Res.* **3**, 20–25 (1994).
- Christenson, R. K. & Day, B. N. Maintenance of unilateral pregnancy in the pig with induced corpora lutea. *J. Anim. Sci.* **32**, 282–286 (1971).
- Prather, R. S. *et al.* Artificial activation of porcine oocytes matured in vitro. *Mol. Reprod. Dev.* **28**, 405–409 (1991).
- McGrath, J. & Solter, D. Nuclear transplantation in the mouse embryo by microsurgery and cell fusion. *Science* **220**, 1300–1302 (1993).
- McGrath, J. & Solter, D. Nuclear transplantation in mouse embryos. *J. Exp. Zool.* **228**, 355–362 (1993).
- Kwon, O. Y. & Kono, T. Production of identical sextuplet mice by transferring metaphase nuclei from four-cell embryos. *Proc. Natl Acad. Sci. USA* **93**, 13010–13013 (1996).
- Wells, D. N., Misica, P. M. & Tervit, H. R. Production of cloned calves following nuclear transfer with cultured adult mural granulosa cells. *Biol. Reprod.* **60**, 996–1005 (1999).
- Campbell, K. H., Loi, P., Otaegui, P. J. & Wilmut, I. Cell cycle co-ordination in embryo cloning by nuclear transfer. *Rev. Reprod.* **1**, 40–46 (1996).
- McCreath, K. J. *et al.* Homologous recombination in ovine somatic cells enables the production of gene targeted sheep by nuclear transfer. *Nature* **405**, 1066–1069 (2000).
- Platt, J. L. Xenotransplantation: recent progress & current perspectives. *Curr. Opin. Immunol.* **8**, 721–728 (1996).
- Petters, R. M. & Wells, K. D. Culture of pig embryos. *J. Reprod. Fert. Suppl.* **48**, 61–73 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank B. Gragg, T. Akers, H. Bishop and J. McPherson for technical contributions to embryo production, embryo transfer and animal husbandry; P. Jobst for technical contributions in embryo production and nuclear transfer; W. Eyestone and K. Polson for valuable contributions in discussions; S. Pleasant and M. Crisman and their staff at the Virginia Maryland Regional Veterinary College for their critical roles in the Caesarian delivery and post-natal care of the cloned piglets. This research was funded in part by a grant from the National Institute of Science and Technology (NIST) Advanced Technology Program (ATP).

Correspondence and requests for materials should be addressed to I.P. (e-mail: ipolejaeva@ppl-therapeutics.com).

Infection by porcine endogenous retrovirus after islet xenotransplantation in SCID mice

Luc J.W. van der Laan*, **Christopher Lockey†**, **Bradley C. Griffeth***, **Francine S. Frasier***, **Carolyn A. Wilson‡**, **David E. Onions§**, **Bernhard J. Hering ||**, **Zhifeng Long†**, **Edward Otto†**, **Bruce E. Torbett*** & **Daniel R. Salomon***

*The Scripps Research Institute, Department of Molecular and Experimental Medicine, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

† Genetic Therapy Inc., A Novartis Company, 9 West Watkins Mill Road, Gaithersburg, Maryland 20878, USA

‡ Food and Drug Administration, Centre for Biologics Evaluation and Research, 8800 Rockville Pike, Bethesda, Maryland 20892, USA

§ University of Glasgow, Department of Veterinary Pathology, Bearsden Road, Glasgow G61 1QH, Scotland & Q-One Biotech Ltd, Todd Campus, Glasgow G20 0XA, UK

|| University of Minnesota, Department of Surgery, 420 Delaware Street S.E., Minneapolis, Minnesota 55455, USA

Animal donors such as pigs could provide an alternative source of organs for transplantation. However, the promise of xenotransplantation is offset by the possible public health risk of a cross-species infection^{1,2}. All pigs contain several copies of porcine endogenous retroviruses (PERV)^{3,4}, and at least three variants of PERV can infect human cell lines *in vitro* in co-culture, infectivity and pseudotyping experiments^{3,5–7}. Thus, if xenotransplantation of pig tissues results in PERV viral replication, there is a risk of spreading and adaptation of this retrovirus to the human host. C-type retroviruses related to PERV are associated with malignancies of haematopoietic lineage cells in their natural hosts⁸. Here we show that pig pancreatic islets produce PERV and can infect human cells in culture. After transplantation into NOD/SCID (non-obese diabetic, severe combined immunodeficiency) mice, we detect ongoing viral expression and several tissue compartments become infected. This is the first evidence that PERV is transcriptionally active and infectious cross-species *in vivo* after transplantation of pig tissues. These results show that a concern for PERV infection risk associated with pig islet xenotransplantation in immunosuppressed human patients may be justified.

Juvenile-onset diabetes mellitus is a major health problem and exogenous insulin therapy is only partially successful in preventing its many complications. Although islet transplantation holds great promise for a cure, the number of potential human pancreas donors are extremely unlikely to provide enough islet tissue to treat the millions of patients worldwide. The xenotransplantation of pig

islets is therefore being actively investigated as a potential alternative clinical therapy. It is still uncertain whether cellular transplantation has the same infectious risk profile as vascularized organ xenotransplantation. To develop a small-animal model for islet xenotransplantation suitable for studying the risk of PERV infection in immunosuppressed human patients, we chose the NOD/SCID mouse strain. This strain can be successfully engrafted with human haematopoietic cells and tissues^{9–12} as well as pig islets which function to reverse chemically induced diabetes (unpublished data).

We first determined the potential of adult pig islets to produce infectious PERV using *in vitro* co-culture assays. Cultured pig islets express PERV mRNA based on quantitative RT–PCR (polymerase chain reaction after reverse transcription of RNA) for the spliced product of PERV *env* sequence (data not shown). This result is consistent with preliminary data that reverse transcriptase (RT) activity was detected in cultures of fetal pig islets¹³. Next, neomycin-resistant U293 cells were co-cultured with pig islets for eight weeks. The neomycin analogue, G418, was added after seven days to eliminate the pig cells. The use of neomycin selection has the advantage of avoiding irradiation, a procedure that might activate

retroviral transcription. As shown in Fig. 1a, the U293 cells were clearly positive for PERV *pol* DNA at 18 and 28 days of co-culture. However, no pig-specific cytochrome *c* sequence could be detected by Southern blotting of PCR products after 11 days of G418 selection (a total of 18 days of culture), indicating that all the pig cells were eliminated. Infection was confirmed by quantitative PCR (Fig. 1b) demonstrating detectable PERV-specific *pol* sequence and undetectable pig centromeric sequence. We demonstrated that U293 cells infected by pig islet-derived PERV secrete infectious virions by showing reverse transcriptase activity and infectivity of these culture supernatants (Fig. 1c). Therefore, purified pig islets release infectious PERV virions and productively infect the human epithelial cell line, U293. U293 is a well established target cell line for retrovirus infection; however, we and others⁶ have also infected mouse cell lines with PERV in culture.

Pig islets transplanted under the kidney capsules of immunodeficient mice were explanted 7–8 weeks later and analysed by con-

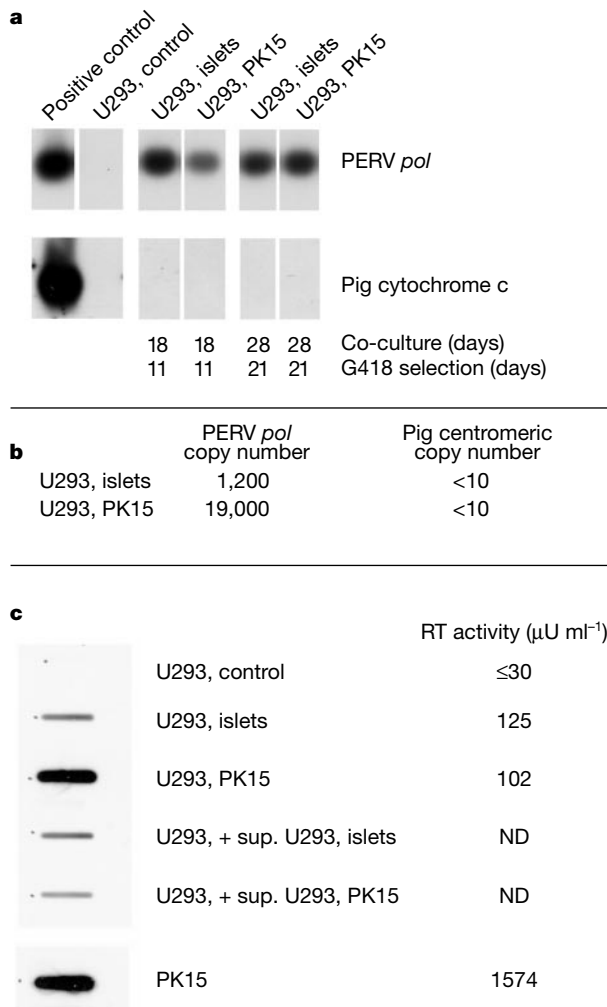
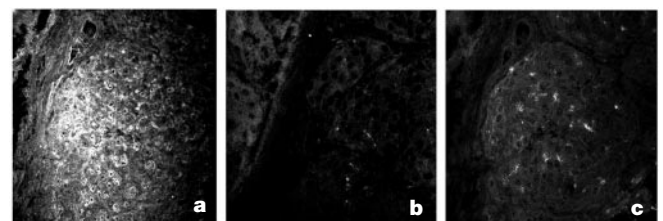


Figure 1 Pig islets release infectious PERV virions and productively infect human U293 cells. **a**, U293/neo cells co-cultured with pig islets or PK15 became clearly positive for PERV *pol* whereas no pig-specific cytochrome *c* sequence could be detected after G418 selection (one of four similar experiments is shown). **b**, The number of PERV *pol* and pig centromeric DNA copies per 100 ng DNA from the same cells after 44 days of culture was quantified ($n = 2$). **c**, Cells incubated with conditioned media from PERV-infected cells clearly became positive, demonstrating that U293 co-cultured with islets are productively infected. RT enzyme activity was tested at week 6 of culture. ND, not done.



d PERV *env* mRNA expression by pig islets after transplantation

	Number of islets	Transplant site	Days after transplantation	Number of PERV positive islet grafts (%)	PERV cDNA copy number per μ g RNA
I	5,000	Kidney capsule	18	2/2 (100)	2.8; 10
II	2,000	Kidney capsule	49	1/3 (33)	7.7
	2,000	Perispinal template	49	0/3 (0)	N/A
III	3,000	Kidney capsule	56	1/3 (33)	28
	3,000	Perispinal template	56	1/3 (33)	5
IV	8,000	Kidney capsule	49	2/2 (100)	9; 1,300
	8,000	Perispinal template	49	3/3 (100)	110; 6,700; 10,000

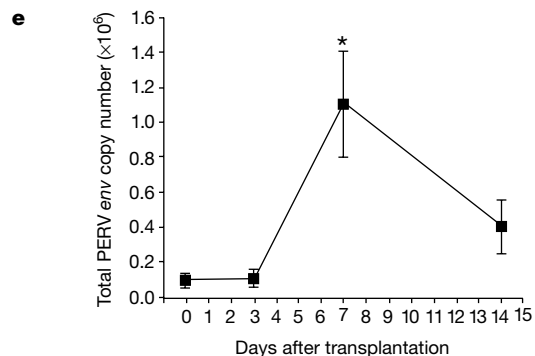


Figure 2 PERV expression in pig islets after transplantation. **a, c**, Immunohistochemical staining for PERV Gag protein p30 (**c**) and insulin (**a**) in pig islets transplanted under the mouse kidney capsule. A number of clearly p30-positive islet cells are identified. **b**, The rabbit antisera control. Magnification, $\times 400$. **d**, Spliced mRNA for PERV *env* was determined by quantitative RT–PCR and detected in islet grafts after transplantation in NOD/SCID mice. N/A, not applicable. **e**, PERV *env* mRNA expression in islet grafts was significantly increased at day 7 (asterisk, $p < 0.015$, Wilcoxon) compared to day 0 and day 3 after transplantation. Shown are the means $\pm$ s.e.m. ($n = 9$) of three independent experiments.

Table 1 Pig/mouse chimaerism in NOD/SCID after pig islet transplantation

(a)							
Experiment	Number of islets	Transplant site	Weeks after transplantation	Number of chimaeric animals	Number of positive tissues chimaeric animals		
I	5,000	Kidney capsule	2.5	3/3	5/6		
II	2,000	Kidney capsule	7	2/4	5/10		
	2,000	Perispinal template	7	3/4	3/14		
III	3,000	Kidney capsule	8	3/3	5/21		
	3,000	Perispinal template	8	2/3	2/14		
IV	8,000	Kidney capsule	7	2/2	4/10		
	8,000	Perispinal template	7	1/4	1/5		
V	(6 × 8,000)	Perispinal template	8	1/2	1/6		
Total				17/25 (68.0%)	26/86 (30.2%)		
.....							
(b)							
Tissues	Spleen	Liver	Salivary gland	Skin	Small bowel	Lung	Kidney
Chimaerism	7/25	6/25	3/22	3/10	4/12	3/14	0/18
Percentage	28	24	14	30	33	21	0

Twenty-five mice in five independent experiments were transplanted with purified pig islets. After 2.5 to 8 weeks animals were killed and various tissues were removed and analysed by quantitative PCR. (a), Seventeen of the twenty-five mice (68%) had one or more tissues other than the transplantation site, that tested positive for both PERV and centromeric DNA. This indicates a state of microchimaerism with pig cells originating from the islet graft. A total of 126 tissue samples were analysed and 30% of the 86 tissues from chimaeric animals were found to contain pig DNA. Many mouse tissues were negative, showing that our primers are not detecting murine endogenous retroviral sequences. (b), incidence of chimaerism for specific tissues.

focal immunohistochemistry for expression of PERV p30 Gag protein. Several cells within the islet grafts were stained positive for PERV p30 (Fig. 2c). This excludes the possibility that p30 expression is restricted to a small population of passenger leukocytes or islet dendritic cells. Moreover, insulin staining (Fig. 2a) did not specifically co-localize with p30 staining. Staining was characteristically in a granular pattern and essentially identical to the staining that we have documented in U293 and primary human bone marrow endothelial cells infected with PERV in co-cultures. At various time points after transplantation (18 to 56 days), islets under the kidney capsule or in three-dimensional collagen templates coated with an arginine, glycine and aspartic acid (RGD) peptide placed in perispinal muscle were analysed for the presence of spliced PERV *env* mRNA using quantitative RT-PCR (Fig. 2d). PERV mRNA was detected in islet grafts in four independent experiments. Inflammatory cytokines and local tissue ischaemia associated with surgical injury might increase the patient's exposure to PERV by activating viral replication shortly after transplantation. We determined PERV mRNA expression in the islets before transplantation and from islets explanted from the same animals 3, 7 and 14 days after transplantation. As shown in Fig. 2e, there is a greater than tenfold increase in subgenomic PERV mRNA expression from

day 0 to day 7. Transcriptional activation of PERV genes thus occurs *in vivo* after pig islet transplantation.

Retrospective analysis of patients in clinical trials exposed *in vivo* or extracorporeally to pig cells has demonstrated long-term survival or microchimaerism of pig cells in some patients. Chimaerism after pig islet transplantation was detected by PCR for mitochondrial DNA in the peripheral blood at six months in four of ten patients¹³. Chimaerism was detected by PCR for pig mitochondrial and centromeric sequences in 23 of 100 patients after extracorporeal splenic perfusion, persisting for up to 8.5 years in one patient¹⁴. We determined the extent of chimaerism in NOD/SCID mice transplanted with pig islets. In 17 of 25 mice (68%), one or more tissues other than the transplanted islets tested positive for PERV *pol* and centromeric DNA (Table 1). In the chimaeric animals, 30% of tested tissues contained pig sequences. Thus, our results show the migration of pig cells after islet transplantation into several tissue compartments increasing the range of tissue exposure to potential PERV infection far beyond the immediate transplantation site.

It is vital to determine whether pig islets release infectious virus after transplantation. Every pig cell contains multiple copies of PERV proviral DNA^{3,4} and centromeric sequence repeats. Our approach to distinguish pig cell chimaerism from a true infection

Table 2 Evidence of PERV infection in tissues of pig/mouse chimaeric mice

Experiment	ID number		PERV copy number		CENT copy number		P/C ratio
			Average ± s.d.	n	Average ± s.d.	n	
I	919	Spleen	145000	2	2900000 ± 173205	3	0.050
		Liver	47 ± 11	3	197 ± 12	3	0.237
	920	Islet/kidney	14000	1	110000 ± 30000	3	0.127
		Spleen	320	1	1700 ± 954	3	0.188
II	921	Spleen	350 ± 123	3	3733 ± 306	3	0.094
		Islet/kidney	10670 ± 191	3	85800 ± 11898	3	0.124
	68	Spleen	186 ± 70	3	649 ± 248	3	0.286
		Small bowel	34 ± 13	3	209 ± 71	3	0.162
	926	Islet/kidney	47300 ± 8305	3	506000 ± 19053	3	0.093
		Small bowel	924 ± 375	3	2156 ± 331	3	0.429
IV	928	Islet/kidney	14300 ± 2807	3	144100 ± 13739	3	0.099
	486	Islet/kidney	506 ± 69	3	2728 ± 529	3	0.186
V	449	Small bowel	391 ± 109	16	444 ± 228	19	0.881
Total chimaeric mice			(n = 8)				0.252 ± 0.260*
Pig tissue controls			(n = 52)				0.043 ± 0.040
Total islet/kidney samples			(n = 5)				0.126 ± 0.036†§
Pig islet preparation controls			(n = 5)				0.050 ± 0.009

Data are based on the pig PERV and centromeric DNA (P/C) ratios. Average and standard deviation (s.d.) of PERV and centromeric DNA copy numbers per 3.3 µg DNA in all 13 chimaeric mouse tissues for which an accurate P/C ratio could be calculated are shown here. These 13 chimaeric tissues were found in 8 mice. We note that in Table 1, there were actually 26 total tissues in 17 animals positive for PERV and pig centromeric DNA. However, 13/26 tissues in 9 animals had PERV values between 1 and 10 copies per 3.3 µg DNA, which, although positive by our assay, are too low for confidence in calculating an accurate P/C ratio and were not analysed further. The average P/C ratio of the 8 remaining chimaeric mice together are significantly different from the ratios of pig tissue controls from 52 individual pigs (**p* < 0.001, Wilcoxon) and the ratios of islet preparations used in these transplants from 5 different pigs (†*p* = 0.002). Also the P/C ratios from the five individual mouse kidneys with islet grafts are significantly different from the ratios of the five control islet preparations (§*p* = 0.008) and the pig tissue controls (§*p* = 0.001). These results indicate that at least 8 of 25 animals (32%) were infected by PERV in at least one tissue compartment after pig islet xenotransplantation.

of mouse tissues is based on the premise that the ratio between PERV proviral DNA and pig centromeric sequences (*P/C* ratio) in any given pig cell or pig tissue is fixed. We tested five different islet preparations from outbred Landrace pigs used for these transplants to obtain a mean *P/C* ratio of 0.050 ± 0.009 (s.d.). We also tested liver tissues from 43 Landrace pigs (*P/C* ratio = 0.042 ± 0.041) and peripheral blood cells from 9 Yorkshire/Hampshire pigs (*P/C* ratio = 0.050 ± 0.033) to obtain a mean *P/C* ratio for control pig tissues of 0.043 ± 0.040 (s.d.). Thus, there is consistency between the *P/C* values calculated for islets, liver and peripheral blood cells supporting the premise that the ratio between these two genes is fixed. If the *P/C* ratio is fixed, then a significant increase in the *P/C* ratio in chimaeric tissues can only be achieved by a relative increase in PERV DNA consistent with infection of the mouse tissue. To test the sensitivity of this approach we calculated *P/C* ratios by mixing 20 pig cells with 10^7 human cells (1 pig cell per 5×10^5 human cells) and spiking the mixture with plasmid PERV DNA from 0 to 3×10^4 molecules. The *P/C* ratio was found to be significantly different when the equivalent of more than 1,000 human cells ($>0.01\%$) are infected by a single copy of PERV.

Table 2 shows the actual copy numbers of PERV and pig centromeric DNA in all 13 tissues from 8 animals in which a *P/C* ratio could be calculated. In 9 of the 17 chimaeric animals shown in Table 1, an accurate *P/C* ratio could not be confidently determined, despite positive tests for PERV and pig centromeric DNA, because the PERV values were too low (1–10 copies per 3.3 μ g DNA). The remaining 8 of 25 total animals that demonstrated no evidence of pig cell chimaerism are not shown and were all negative for PERV. The results demonstrate PERV infection in one or more tissues in the 8 chimaeric animals (32%) based on significantly higher *P/C* ratios when compared against the ratios of the 5 pig islet preparations used in these transplantation studies ($P = 0.002$, Wilcoxon) or 52 control pig tissues ($P < 0.001$). A possible complication for our model is that mice are known to contain xenotropic endogenous retroviruses^{15,16}. Thus, it is possible that endogenous murine retroviruses could infect transplanted pig cells and then rescue or pseudotype PERV to allow infection of mouse cells. In addition, recombination of PERV and endogenous murine retroviruses could have resulted in a recombinant PERV virus with altered tropism.

Here we show that pig islets release infectious PERV virus which can productively infect human target cells *in vitro*. After islet transplantation we demonstrate active proviral DNA transcription by measuring subgenomic (spliced) mRNA for the PERV *env* gene and expression of p30 Gag by islet cells. Moreover, we demonstrate PERV infection of mouse cells in multiple tissue compartments. We also demonstrate an extensive chimaerism of pig cells in these animals after transplantation, indicating that the host's exposure to PERV infection will extend well outside the transplantation site. In fact, we have only found infection in compartments chimaeric for pig cells, suggesting that cell–cell contact may be one important mechanism for spreading infection.

It is interesting to consider the recent evidence that PERV infection was not demonstrated in humans exposed to pig cells^{13,14,17}. First, PERV infection in these human studies was only determined with peripheral blood lymphocytes and serum whereas our studies involved several tissue compartments in the mouse. There is evidence suggesting that primary human peripheral blood cells cannot be productively infected with PERV *in vitro*^{3,18}. Furthermore, comparisons of host ranges for PERV consistently demonstrate increased susceptibility for human epithelial cell lines as compared to both T and B cell lines⁶. Secondly, most patients in the clinical series were not immunosuppressed at the time of pig cell exposure, and the majority of immunosuppressed recipients of islets did not demonstrate long-term survival and function of the grafts. Moreover, the retrospective nature of these studies did not allow consecutive cases to be analysed at set time points post-exposure so the possibility of missing evidence of infection in the

peripheral blood cannot be excluded. For example, the 14 patients with pig islet transplants had early post-transplant sera tested for antibodies but peripheral blood lymphocytes were only tested at 4 to 7 years afterwards¹³. Nor can we exclude the possibilities that mouse tissue is more susceptible to infection than human or that pseudotyping of PERV with mouse endogenous retrovirus did not increase the host range in our model.

This study provides evidence of cross-species transmission of PERV *in vivo* following pig tissue transplantation. We have created a model to study PERV infection in an immunodeficient mouse that is permissive for transplantation of both human and pig tissues. A human patient is unlikely to have an immunodeficiency of equivalent magnitude and the mice do not have natural immunity against galactose $\alpha(1-3)$ -galactose terminal sugars^{3,19}. But pig cell chimaerism was detected in human patients, including 23 patients that were not immunosuppressed¹⁴. Moreover, successful xenotransplantation will involve overcoming the natural immune barrier by clinical strategies, such as engineering pigs with complementary regulatory proteins or alterations in enzymatic pathways responsible for sugar modification of cell-surface molecules which will also eliminate this primary immune barrier to PERV infection^{19,20}. Thus, we conclude that successful pig islet xenotransplantation to human patients may result in long-term exposure to replication competent endogenous retrovirus. Our model supports the concern that such an exposure will be associated with PERV infection of cells in multiple tissue compartments. □

Methods

Pig NOD/SCID chimaeras

NODIt/szSCID mice (L. D. Shultz) are maintained under pathogen-free conditions. 2,000–8,000 pig islet equivalents (IEQ) were transplanted under the kidney capsule²¹ or perispinus muscles in three-dimensional collagen templates coated with RGD peptide analogues. These templates were developed with Integra Life Sciences (San Diego) to protect islet structure and enhance angiogenesis. In some experiments (Fig. 2e) multiple templates (8,000 IEQ each) were placed in the perispinus muscles and removed at 3, 7 and 14 days after transplantation for quantification of PERV mRNA.

Cell lines, pig islets and cell co-cultures

U293 (human kidney epithelial, #CRL-1573) and PK15 (pig kidney, #CCL-33) cell lines were obtained from American Type Culture Collection. U293 transfected with a neomycin resistance gene (U293/neo) were selected with $800 \mu\text{g ml}^{-1}$ G418 (Calbiochem). U293 (2×10^4) were co-cultured with either PK15 (2×10^4) or pig islets (2,000 IEQ) for 8 weeks. Infections of U293 by supernatants were done with conditioned media from PERV-infected U293 (passage 10 or later), added in a 2:1 volume with fresh media every 3 days for 12 days after which cells were cultured for 4 weeks. Pig islets were purified from female Landrace pigs. Freshly harvested pancreata underwent intraductal distention with solution containing 0.27–0.33% of collagenase (Serva). Islets were isolated within one hour by a continuous digestion–filtration device and purified on a discontinuous gradient (Biochrom)²².

Immunohistochemistry

Detection of p30 PERV-expression was done with rabbit polyclonal anti-recombinant p30-specific antisera and donkey Fab₂ FITC anti-rabbit IgG (Jackson Labs). Similar results were obtained with rabbit anti-p30 peptide-specific antisera. Insulin was stained with sheep anti-human insulin (The Binding Site) followed by donkey Fab₂ LSRC anti-sheep IgG (Jackson Labs). Sections were blocked with donkey serum and secondary reagents were pre-absorbed to eliminate cross-species reactivity. Sections were analysed with a Zeiss microscope equipped with a scanning laser confocal attachment (MRC 1024, BioRad).

RT assay, PCR, RT–PCR and Southern blotting

RT activity was measured in cell supernatants using a Mn^{2+} -dependent RT assay (Cavidi HS-kit). Supernatants after 7 days of culture were diluted 1:15 and 1:30 and incubated overnight with an immobilized template/primer and BrdUTP. RT activity was measured by BrdUMP incorporation using a specific alkaline phosphatase-conjugated antibody. Results were compared to MMuLV rRT and expressed as $\mu\text{U ml}^{-1}$. Values less than two times the background ($<30 \mu\text{U ml}^{-1}$) were below the limit of detection.

PERV sequences were detected by PCR with conserved *pol* sequence-derived probes¹⁴: 5'-AGCTCCGGGAGGCGCTACTC-3' and 5'-ACAGCCGTTGGTGTGGTCA-3'. Pig-specific cytochrome *c* DNA was detected with 5'-CATTCTACGAGGTCTGTTCCG-3' and 5'-GCCTATTCATCCACGTAGGC-3' probes (GenBank accession no. U18827). Cycling parameters were 45 s at 64 °C, 60 s at 72 °C and 30 s at 94 °C for 35 cycles. A *pol* gene-containing plasmid was used as a positive control for PERV, and PK15 DNA as a positive control for cytochrome *c*. Detection of PCR products was done by Southern or slot blotting (Schleicher & Schuell) of DNA using digoxigenin-labelled probes (full-length

PERV *pol* and pig cytochrome *c* PCR products) and rabbit anti-digoxigenin alkaline phosphatase (Boehringer Mannheim). Quantitation of PERV sequences using fluorescence-based, real-time PCR (ABI/Perkin-Elmer) was performed in 100 μ l containing 5U Taq polymerase, 1U UNG, 300 nM of each primer and 100 nM of probe 5'-FAM-CCACCGTGCAGGAACCTCGAGACT-TAMRA-3'. Cycling parameters were 2 min at 50 °C, 10 min at 95 °C and then 60 cycles of 15 s at 95 °C and 1.5 min at 60 °C. The limit of detection for this assay is one PERV copy in 3.3 μ g genomic DNA (500,000 cells); the sensitivity of the assay was defined as 10 copies per 500,000 cells, which gave >99.99% confidence of detecting ≥ 10 copies¹⁴.

DNA samples were tested for the presence of porcine centromeric sequences using primers 5'-TAGCCATGCTGCATGTAATGC-3', 5'-GGAGCGTGGCCCAAT-3' and probe 5'-FAM-ATGCTGCATGGAATGCACTACCTTCAA-TAMRA-3'¹⁴. Less than 10 copies in 3.3 μ g DNA was considered negative. The number of PERV and centromeric sequence repeats in 43 pig liver samples was 67 ± 7.2 (s.e.m.) and $3,694 \pm 688$ per cell, respectively. Because the number of PERV copies per cell did not conform to a gaussian distribution, the *P/C* ratios were determined for each of the 43 control samples individually. This approach yielded a mean *P/C* ratio of 0.042 rather than the value of 0.018, which is the simple ratio of the values for the average PERV versus centromeric copy numbers (67/3700). Significance of detected differences was tested by Wilcoxon Two Sample test (<http://fons3.let.uva.nl/Service/Statistics.html>).

We determined the level of a spliced form of the PERV *env* mRNA using primers to a conserved region. Total RNA was isolated using RNeasy (Qiagen) and reverse transcribed by priming with oligo d(T)₁₆ and 75U MuLV RT (10 min at 25 °C followed by 15 min at 42 °C, then 5 min at 99 °C). cDNA was amplified using primers and probe specific for PERV *env*: 5'-GTTTGCATCAAGACCGCTTCT-3' (300 nM) and 5'-AGCCGTGTGTGTGTCAGAAA-3' (300 nM) and 5'-FAM-CCACCGTGCAG-GAAACCTCGAGACT-TAMRA-3' (75 nM) (GenBank accession no. AF038601). Cycling parameters were 2 min at 50 °C, 10 min at 95 °C and then 50 cycles of 15 s at 95 °C and 1 min at 60 °C. A standard curve of 10^2 to 10^6 copies of plasmid containing spliced *env* cDNA allowed quantitation of starting cDNA concentrations.

Received 5 April; accepted 22 June 2000.

- Chapman, L. E. *et al.* Xenotransplantation and xenogeneic infections. *N. Eng. J. Med.* **333**, 1498–1501 (1995).
- Stoye, J. P. Xenotransplantation. Proviruses pose potential problems. *Nature* **386**, 126–127 (1997).
- Patience, C., Takeuchi, Y. & Weiss, R. A. Infection of human cells by an endogenous retrovirus of pigs. *Nature Med.* **3**, 282–286 (1997).
- Akiyoshi, D. E. *et al.* Identification of a full-length cDNA for an endogenous retrovirus of miniature swine. *J. Virol.* **72**, 4503–4507 (1998).
- Le Tissier, P., Stoye, J. P., Takeuchi, Y., Patience, C. & Weiss, R. A. Two sets of human-tropic pig retrovirus. *Nature* **389**, 681–682 (1997).
- Takeuchi, Y. *et al.* Host range and interference studies of three classes of pig endogenous retrovirus. *J. Virol.* **72**, 9986–9991 (1998).
- Wilson, C. A. *et al.* Type C retrovirus released from porcine primary peripheral blood mononuclear cells infects human cells. *J. Virol.* **72**, 3082–3087 (1998).
- Toth, S. R., Onions, D. E. & Jarrett, O. Histopathological and hematological findings in myeloid leukemia induced by a new feline leukemia virus isolate. *Vet. Pathol.* **23**, 462–470 (1986).
- Hesselton, R. M. *et al.* High levels of human peripheral blood mononuclear cell engraftment and enhanced susceptibility to human immunodeficiency virus type 1 infection in NOD/LtSz-scid/scid mice. *J. Infect. Dis.* **172**, 974–982 (1995).
- McCune, J. M. *et al.* The SCID-hu mouse: murine model for the analysis of human hematolymphoid differentiation and function. *Science* **241**, 1632–1639 (1988).
- Crisa, L. *et al.* Human cord blood progenitors sustain thymic T-cell development and a novel form of angiogenesis. *Blood* **94**, 3928–3940 (1999).
- Miyoshi, H., Smith, K. A., Mosier, D. E., Verma, I. M. & Torbett, B. E. Transduction of human CD34+ cells that mediate long-term engraftment of NOD/SCID mice by HIV vectors. *Science* **283**, 682–686 (1999).
- Heneine, W. *et al.* No evidence of infection with porcine endogenous retrovirus in recipients of porcine islet-cell xenografts. *Lancet* **352**, 695–699 (1998); erratum *Lancet* **352**, 1478 (1998).
- Paradis, K. *et al.* Search for cross-species transmission of porcine endogenous retrovirus in patients treated with living pig tissue. *Science* **285**, 1236–1241 (1999).
- Levy, J. A. Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB, and other mouse strains. *Science* **182**, 1151–1153 (1973).
- Lamont, C. *et al.* Characterization of endogenous and recombinant proviral elements of a highly tumorigenic AKR cell line. *J. Virol.* **65**, 4619–4628 (1991).
- Patience, C. *et al.* No evidence of pig DNA or retroviral infection in patients with short-term extracorporeal connection to pig kidneys. *Lancet* **352**, 699–701 (1998).
- Wilson, C. A., Wong, S., VanBrocklin, M. & Federspiel, M. J. Extended analysis of the in vitro tropism of porcine endogenous retrovirus. *J. Virol.* **74**, 49–56 (2000).
- Sandrin, M. S. & McKenzie, I. F. Recent advances in xenotransplantation. *Curr. Opin. Immunol.* **11**, 527–531 (1999).
- Weiss, R. A. Transgenic pigs and virus adaptation. *Nature* **391**, 327–328 (1998).
- Beattie, G. M. & Hayek, A. Outcome of human fetal pancreatic transplants according to implantation site. *Transpl. Proc.* **26**, 3299 (1994).
- Liu, X. *et al.* Prolongation of porcine islet allograft survival. *Transpl. Proc.* **29**, 768–769 (1997).

Acknowledgements

We thank J. Elder, A. Pelletier, J. Coffin, J. Stoye and G. Langford for helpful discussions, and D. Stablein for expert statistical advice. We thank I. Mychkovsky and T. Gladden for assistance in mRNA/DNA purification, T. Gilmore, J. Ansie and D. Scharpe for assistance in purification of pig islets, and L. Crisa, P. Hildbrand and O. Schussler for help with mouse surgery. We thank R. Ingram and M. Pierschbacher for the RGD collagen templates

for transplantation. D.R.S., B.C.G., F.S.F., B.E.T. and this work are supported by an NIH/NIAD grant. L.J.W.L. is supported by a Juvenile Diabetes Foundation International (JDFI) fellowship.

Correspondence and requests for materials should be addressed to: D.R.S. (email: dsalomon@scripps.edu).

AT₁-receptor heterodimers show enhanced G-protein activation and altered receptor sequestration

Said AbdAlla*, Heinz Lothar† & Ursula Quittner‡

* Genetics Engineering and Biotechnology Research Institute, Manhiat El-Olama, El-Dekheila, Alexandria, Egypt

† Heinrich Pette Institut, Martinstraße 52, 20251 Hamburg, Germany

‡ Institut für Pharmakologie, Versbacher Straße 9, 97078 Würzburg, Germany

The vasopressor angiotensin II regulates vascular contractility and blood pressure through binding to type 1 angiotensin II receptors (AT₁; refs 1, 2). Bradykinin, a vasodepressor, is a functional antagonist of angiotensin II (ref. 3). The two hormone systems are interconnected by the angiotensin-converting enzyme, which releases angiotensin II from its precursor and inactivates the vasodepressor bradykinin⁴. Here we show that the AT₁ receptor and the bradykinin (B₂) receptor also communicate directly with each other. They form stable heterodimers, causing increased activation of G α_q and G α_i proteins, the two major signalling proteins triggered by AT₁. Furthermore, the endocytotic pathway of both receptors changed with heterodimerization. This is the first example of signal enhancement triggered by heterodimerization of two different vasoactive hormone receptors.

To identify proteins interacting with the active form of the B₂ receptor, we developed a protocol of ligand affinity-labelling^{5,6} before immuno-affinity chromatography. This enabled us selectively to detect proteins interacting with active agonist-bound receptors. We crosslinked bradykinin to the B₂ receptor and performed immuno-affinity chromatography with anti-B₂-receptor antibodies. B₂-receptor monomers (of relative molecular mass, 60,000; *M_r* 60K) and dimers (*M_r* 115K) were enriched as detected by antibodies to bradykinin or to the B₂ receptor (Fig. 1a). The AT₁ receptor was co-enriched by the anti-B₂-receptor antibodies. The B₂-specific antagonist Icatibant abolished the labelling of the B₂ receptor by bradykinin, but did not change the number of detected receptor dimers.

To confirm the identity of the co-enriched protein, we labelled the AT₁ receptor (*M_r* 55K) with ¹²⁵I-labelled angiotensin II (Fig. 1b). Anti-B₂-receptor antibodies enriched ¹²⁵I-labelled AT₁-receptor dimers but AT₁-receptor monomers were not enriched. Unrelated antibodies did not enrich the labelled receptors and labelling of AT₁ was suppressed by a 200-fold molar excess of unlabelled angiotensin II. Neither B₂-specific bradykinin nor PD123319, an AT₂-specific ligand, suppressed the labelling by ¹²⁵I-labelled angiotensin II (not shown). Together, these results confirm that the protein labelled by ¹²⁵I-labelled angiotensin II is AT₁. Reciprocally, after crosslinking of angiotensin II to the AT₁ receptor, anti-AT₁-receptor antibodies enriched a protein complex of AT₁ and B₂ receptors (see Supplementary Information). These results indicate that active high-affinity AT₁ and B₂ receptors form a heteromeric complex on smooth muscle cells.

Under which conditions do AT₁ and B₂ receptors form heterodimers? In HEK-293 cells expressing AT₁ (ref. 7) or B₂ (ref. 8) receptors, the monomeric form was predominant (Fig. 2a and Supplementary Information). The anti-AT₁-receptor antibodies did not crossreact with B₂ or vice versa. In cells co-expressing AT₁ and B₂, a dimeric receptor form was detected by both anti-receptor antisera. The dimeric form consists of AT₁/B₂-receptor heterodimers, as AT₁ receptors of high molecular mass were only co-enriched by anti-B₂-receptor antibodies (or vice versa) from cells expressing both AT₁ and B₂ receptors, but not from cells expressing either B₂ or AT₁ receptors alone (Fig. 2a). AT₁/B₂-receptor heterodimers were stable under non-reducing conditions of SDS-polyacrylamide gel electrophoresis (PAGE) without prior crosslinking and persisted in the absence of agonist. As such stable dimers were not identified under conditions of agonist-induced homodimerization⁶, the biochemical characteristics of AT₁/B₂-receptor heterodimers are different from B₂-receptor homodimers. The AT₁/B₂ heterodimers were sensitive to reducing agents (Fig. 2b, c).

To analyse whether the presence of the second receptor is sufficient for AT₁/B₂-receptor heterodimerization, the number of B₂ receptors was increased in AT₁-receptor-expressing cells and vice versa. Increasing the number of B₂ receptors in AT₁-expressing cells produced an angiotensin-II-reactive protein complex which was absent in cells expressing AT₁ receptors only (Fig. 2d). We found similar results for the B₂ receptor after affinity-labelling by bradykinin (not shown) or after labelling by the B₂-specific antagonist Icatibant when the expression of AT₁ was increased in B₂-receptor-expressing cells (Fig. 2e). Specificity of the affinity-labelling and of the detection of AT₁ and B₂ receptors by the anti-ligand antibodies is presented as Supplementary Information. These results demonstrate that AT₁/B₂-receptor heterodimer formation is not agonist-

dependent, and that the number of B₂ and AT₁ receptors determines the extent of heterodimerization.

We next analysed the functional properties of AT₁/B₂-receptor heterodimers. Agonist-induced B₂-receptor homodimerization is involved in receptor sequestration and desensitization⁶. In cells expressing AT₁ or B₂ receptors only, the extent of AT₁ or B₂ receptor sequestration was similar to that of cells co-expressing AT₁ and B₂ receptors (Fig. 3a). However, the endocytotic pathway of both receptors changed. AT₁- or B₂-receptor sequestration in HEK-293 cells expressing AT₁ or B₂ receptors involves a dynamin- and clathrin-independent pathway⁹ and therefore was insensitive to a dominant-negative dynamin K44A mutant (Fig. 3a). By contrast, angiotensin II- or bradykinin-induced receptor sequestration in cells co-expressing AT₁ and B₂ was suppressed by the dynamin K44A mutant (Fig. 3a). These findings indicate that upon heterodimerization, AT₁ and B₂ receptors internalize by a dynamin-dependent mechanism.

We examined the receptor/G-protein interaction by following the receptor-stimulated redistribution of G-proteins into the cytosol¹⁰. In cells expressing AT₁ receptors, angiotensin II stimulation did not lead to the detection of significant amounts of Gα in the cytosolic fraction, whereas in cells expressing B₂ receptors cytosolic Gα was

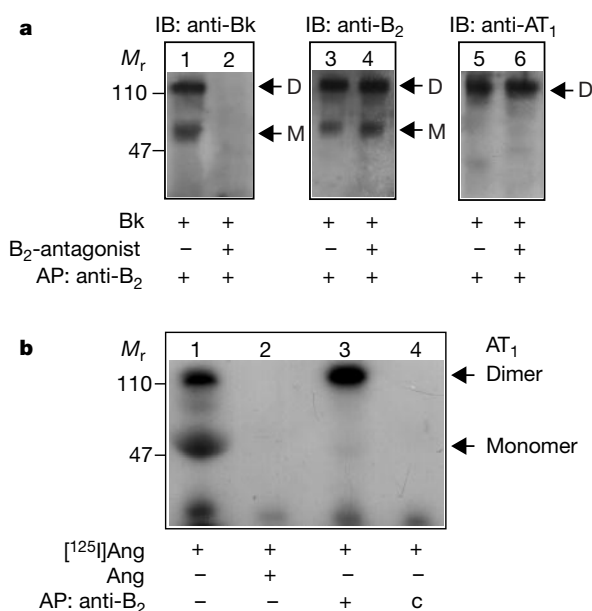


Figure 1 Co-enrichment of AT₁ with B₂ receptors. **a**, Bradykinin (Bk) was crosslinked to B₂ receptors of smooth muscle cells in the absence (lanes 1,3,5) or presence (lanes 2,4,6) of a 200-fold molar excess of the B₂-antagonist Icatibant. Proteins were enriched by immuno-affinity purification (AP) with anti-B₂-receptor antibodies, and enriched receptors were identified in immunoblot (IB) by anti-bradykinin, anti-B₂-receptor or anti-AT₁-receptor antibodies. **b**, Autoradiography shows co-enrichment of [¹²⁵I]-labelled angiotensin II ([¹²⁵I]Ang)-labelled AT₁ receptors and B₂ receptors. [¹²⁵I]Ang was crosslinked to AT₁ in the absence (lanes 1, 3, 4) or presence (lane 2) of a 200-fold molar excess of unlabelled angiotensin II. Immuno-affinity purification (AP) used anti-B₂-receptor antibodies (lane 3) or control antibodies (lane 4). D, dimer; M, monomer.

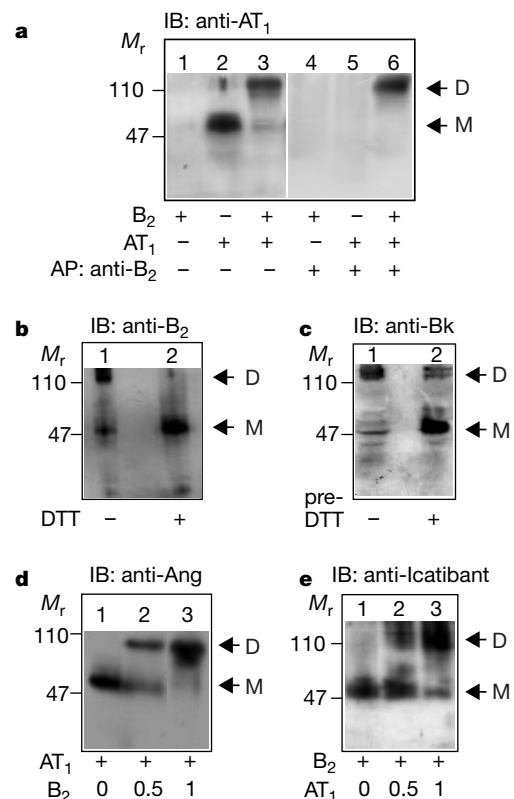


Figure 2 Formation of AT₁/B₂ heterodimers on HEK-293 cells. **a**, Detection of AT₁ receptors by anti-AT₁ antibodies in immunoblot (IB) of cells expressing B₂ (lanes 1, 4), AT₁ (lanes 2, 5) or both (lanes 3, 6) receptors without (lanes 1–3) or with (lanes 4–6) prior immuno-affinity purification (AP) by anti-B₂-receptor antibodies. **b**, Immunoblot detection of B₂ receptors in membranes of cells expressing AT₁ and B₂ receptors by anti-B₂-receptor antibodies after SDS-PAGE performed in the absence (lane 1) or presence (lane 2) of DTT. **c**, Pretreatment of intact cells by DTT (pre-DTT) followed by *N*-ethylmaleimide and washing before affinity-labelling of B₂ receptors with bradykinin (lane 2), suppresses the detection of AT₁/B₂-receptor heterodimers (lane 1) by anti-bradykinin antibodies (anti-Bk). **d**, Immunoblot detection of angiotensin-II-labelled AT₁ receptors by anti-angiotensin II antibodies (anti-Ang) in membranes of cells expressing AT₁ (lane 1), AT₁ and B₂ in a molar ratio of 1:0.5 (lane 2) or 1:1 (lane 3). **e**, Immunoblot of Icatibant-labelled B₂ receptors in membranes of cells expressing B₂ (lane 1), or B₂ and AT₁ in a molar ratio of 1:0.5 (lane 2) or 1:1 (lane 3). D, dimer; M, monomer.

detected after stimulation by bradykinin (Fig. 3b). By contrast, cytosolic G-protein was detected after angiotensin II stimulation in cells co-expressing AT₁ and B₂ receptors although G-protein expression did not change. Similar results were found for the angiotensin-II-stimulated redistribution of Gα_i or Gα_q (Fig. 3c). Thus, co-expression of AT₁ and B₂ receptors increased the extent of

AT₁-stimulated redistribution of Gα-proteins, possibly caused by heterodimerization of AT₁ and B₂ receptors.

We analysed the AT₁-receptor/G-protein coupling further and determined the angiotensin-II-stimulated formation of inositol phosphates. The potency and efficacy of angiotensin II increased with the receptor density (Fig. 3d), as for other G-protein-coupled

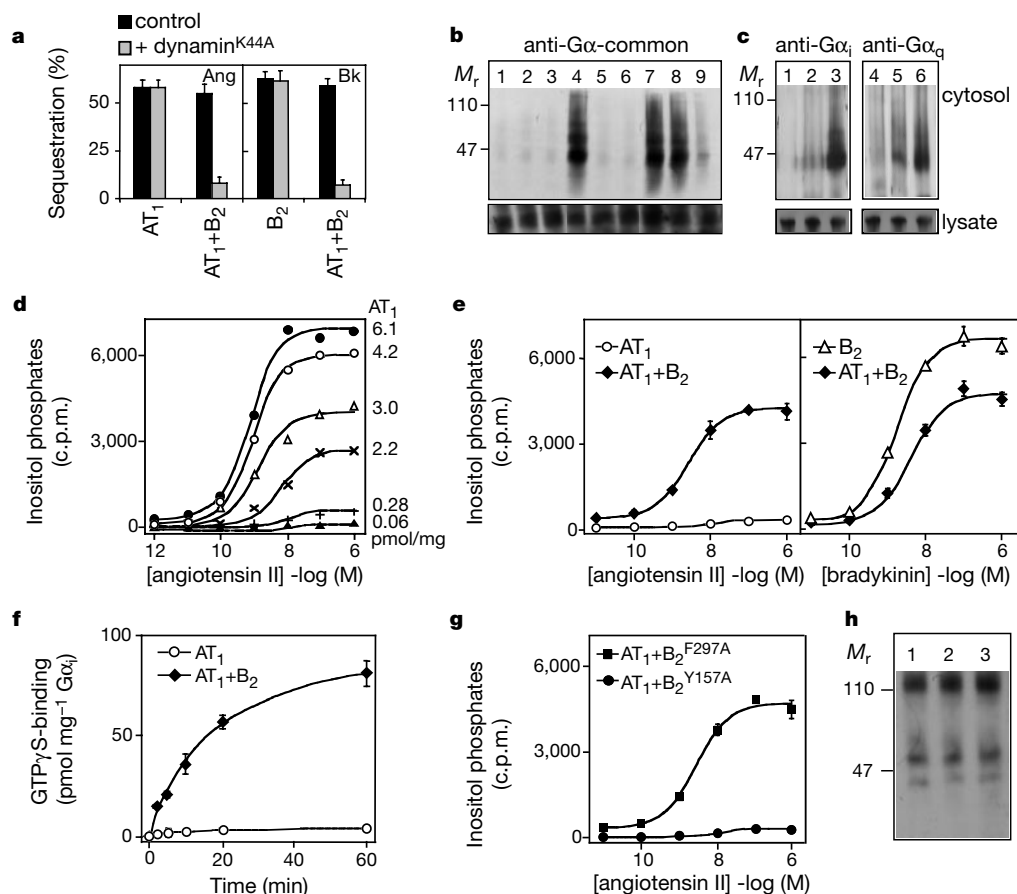


Figure 3 Functional characterization of AT₁–B₂-receptor heterodimers. **a**, Effect of dynamin K44A expression on the angiotensin II (Ang)- or bradykinin (Bk)-stimulated sequestration of AT₁ (left) or B₂ receptors (right). **b**, Agonist-induced redistribution of Gα-proteins in HEK-293 cells expressing AT₁ (lanes 1–3), B₂ (lanes 4–6) or AT₁ and B₂ receptors (lanes 7–9). After addition of bradykinin (lanes 1, 4, 7), angiotensin II (lanes 2, 5, 8) or buffer as a control (lanes 3, 6, 9), Gα-proteins were detected in immunoblot of the cytosolic fractions (top) or of the cell lysates (bottom) by anti-Gα-common antibodies. **c**, Angiotensin-stimulated redistribution of Gα_i (lanes 2, 3) or Gα_q (lanes 5, 6) into the cytosol of cells expressing AT₁ only (lanes 1, 2, 4, 5) or AT₁ and B₂ receptors (lanes 3, 6) determined in immunoblot by anti-Gα_i (lanes 1–3) or anti-Gα_q antibodies (lanes 4–6). In lanes 1 and 4 buffer was used. **d**, Relationship between AT₁-receptor expression and the

angiotensin-II-stimulated increase in inositol phosphates. **e**, Angiotensin-II- (left) or bradykinin- (right) stimulated increase in inositol phosphates in HEK-293 cells expressing the indicated receptors. **f**, Angiotensin-II-stimulated [³⁵S]GTPγS binding to Gα_i. **g**, Angiotensin-II-stimulated increase in inositol phosphates in cells expressing AT₁ and mutated B₂ F297A or AT₁ and B₂ Y157A. Receptor densities in the experiments of e–g were 3,000–10,000 receptors per cell for AT₁ and 10,000–50,000 receptors per cell for B₂ receptors. Shown are the means (± s.e.m.) of 9–11 experiments; (**e**, **g**); three independent experiments (**f**); two independent experiments (**d**). **h**, Immunoblot detection of AT₁/B₂-receptor heterodimers by anti-B₂-receptor antibodies in membranes of cells co-expressing AT₁ and B₂ (lane 1), AT₁ and B₂ F297A (lane 2) or AT₁ and B₂ Y157A (lane 3).

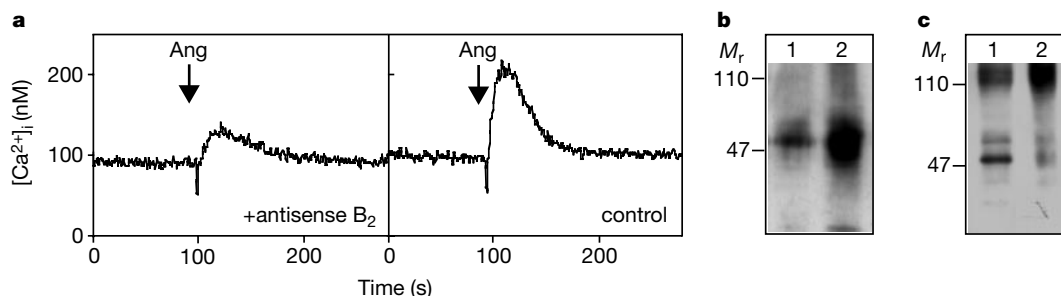


Figure 4 AT₁–B₂-receptor heterodimerization on A10 smooth muscle cells. **a**, Angiotensin-II-stimulated rise in [Ca²⁺]_i in A10 cells after transfection of a B₂ antisense cDNA construct (left) or in control cells (right). **b**, B₂ antisense treatment (lane 1) decreases

the B₂-receptor density compared with control cells (lane 2), as determined in immunoblot by anti-B₂-receptor antibodies. **c**, B₂ antisense treatment (**c**; lane 1) did not reduce the total amount of AT₁ receptors compared with the control (lane 2).

receptors (GPCRs; ref. 11). Subsequent experiments were performed at low AT₁-receptor density (60–100 fmol mg⁻¹) as reported for many *in vivo* systems^{7,12}. Co-expression of AT₁ and B₂ receptors increased the efficacy and potency of angiotensin II (Fig. 3e). By contrast, the potency and the efficacy of bradykinin significantly decreased upon co-expression of the AT₁ receptor. As a control, angiotensin II (10 μM) did not stimulate B₂ receptors, and bradykinin (10 μM) did not stimulate AT₁ receptors (not shown). The number of AT₁ and B₂ receptors as determined by immunoblot (Fig. 2a, b) or by saturation binding did not increase upon co-expression of the second receptor (not shown). Simultaneous stimulation of AT₁/B₂-receptor heterodimers with bradykinin and angiotensin II was additive compared with stimulus alone (not shown). In addition to Gα_q proteins, angiotensin-stimulated activation of Gα_i proteins was increased by heterodimerization of the AT₁ and B₂ receptors, as determined by binding of [³⁵S]GTPγS to Gα_i (Fig. 3f).

The increase in AT₁-receptor-mediated signalling was independent of the binding of bradykinin to the B₂ receptor. Expression of a B₂-receptor mutant with 700-fold decreased affinity for bradykinin¹³, B₂ F297A, also enhanced AT₁-receptor signalling, as did the wild-type B₂ receptor (Fig. 3g). By contrast, co-expression of a B₂-receptor mutant with impaired G-protein coupling¹⁴, B₂ Y157A, did not enhance the AT₁-receptor-mediated signal (Fig. 3g). Both B₂-receptor mutants induced the formation of AT₁/B₂-receptor heterodimers (Fig. 3h). Thus, the signal enhancement by AT₁-receptor heterodimers relies on the intact receptor/G-protein interface of the B₂ receptor.

Finally, we analysed the functional consequences of AT₁/B₂-receptor heterodimerization in native smooth muscle cells. We used a subclone of the smooth muscle cell line A10 expressing AT₁ and B₂ receptors. When the expression of the B₂ receptor decreased by about 80% by an antisense B₂ complementary DNA construct (Fig. 4b), the angiotensin-II-stimulated rise in intracellular Ca²⁺ also decreased by about 70 ± 8% compared with the control (Fig. 4a). By contrast, the amount of AT₁ receptor did not significantly change upon introduction of the B₂ antisense construct (Fig. 4c). However, the intensity of the (hetero)-dimeric form of the AT₁ receptor decreased and the monomeric form increased. These findings indicate that heterodimerization of the AT₁ and B₂ receptors also enhances AT₁-receptor signalling in smooth muscle cells.

The growing list of GPCRs that exist in a dimeric assembly^{6,15–21} implies that dimerization may be a common feature of GPCRs. This work is the first example to our knowledge of signal enhancement triggered by receptor heterodimerization of two vasoactive hormone receptors. AT₁ and B₂ are co-expressed in several tissues, including smooth muscle^{22,23}, glomerular mesangium²⁴ and renal medulla²⁵. As many GPCRs share a common structure, cellular localization and activation mechanism, GPCRs other than AT₁ and B₂ may be structurally compatible to form functional heterodimers. And indeed, κ- and δ-opiate receptors also form heterodimers, thus modifying their signalling properties¹⁶. Thus, heterodimerization between different receptors may define a general mechanism of cell-specific modulation of G-protein-coupled receptors. □

Methods

Cell culture and transfections

HEK-293 and A10 cells were obtained from the American Type Culture Collection, cultured and transfected with the indicated cDNAs as described²⁶.

Receptor-stimulated binding of [³⁵S]GTPγS to Gα

Membranes of HEK-293 cells (0.1 mg protein ml⁻¹) expressing the indicated receptors were prepared as described²⁶, treated by 5 M urea and reconstituted with recombinant His₆-Gα_i, containing an internal hexahistidine-tag at position 121 (40 nM), and Gβγ (60 nM). Receptor-stimulated binding of [³⁵S]GTPγS was determined after the addition of 100 nM angiotensin II as described²⁶.

Functional assays

Inositol phosphate levels of HEK-293 cells were determined as described²⁶. For determination of the intracellular free calcium concentration, [Ca²⁺]_i, of A10 cells, adherent cells

were loaded by fura-2, and changes in [Ca²⁺]_i were determined after stimulation by 100 nM angiotensin II as described²⁷. Receptor sequestration was determined in HEK-293 cells transiently transfected by the indicated cDNAs. Sequestration was defined as the percentage of total cell-surface receptors which, after exposure to agonist, were removed from the plasma membrane⁶.

Affinity-labelling of AT₁ and B₂ and co-enrichment of receptors

Affinity-labelling of AT₁ and B₂ receptors is described in the Supplementary Information. To suppress agonist-induced receptor homodimerization⁶, experiments were performed at 4 °C. Membrane preparation of rat mesenteric smooth muscle cells²², binding of bradykinin or angiotensin II and crosslinking to membranes (0.5–1 mg in 3 ml) was performed as described (see Supplementary Information) using 0.2 mM disuccinimidyl tartarate (DST) instead of 1,5-difluoro-2,4-dinitrobenzene (DFDNB). For crosslinking of [¹²⁵I]-labelled angiotensin II, carrier-free radio-iodinated angiotensin II was used (specific activity 2,000 Ci mmol⁻¹). After washing three times with PBS, membranes were collected by centrifugation, solubilized by 400 μl of 2% (w/v) SDS in 20 mM Tris, pH 7.4, desalted, recovered in 2 ml HGNT buffer⁵ and subjected to immuno-affinity chromatography by anti-B₂-receptor antibodies⁵ or by anti-AT₁-receptor antibodies using 0.1 ml immuno-affinity matrix (Affigel 10; 15 mg affinity-purified antibodies per ml gel). Anti-AT₁-receptor and anti-B₂-receptor antibodies raised against two different epitopes were used for immuno-affinity purification and subsequent immunoblot detection. The antibodies were raised against peptides corresponding to the receptors' amino and carboxy termini. Proteins were eluted by 0.2 M glycine, pH 2.5, neutralized, desalted, delipidated and precipitated²⁸. Proteins were dissolved and separated by 10% SDS-PAGE under reducing conditions, and B₂ and AT₁ receptors were identified in immunoblot by anti-ligand or anti-receptor antibodies.

Receptor-stimulated redistribution of G proteins

Redistribution of G-proteins between cytosol and membranes was determined after agonist stimulation. Transfected HEK-293 cells were incubated by bradykinin or angiotensin II (100 nM) for 30 min at 4 °C. Dithiois (succinimidyl propionate) (2 mM) was added to the cells and the cells were immediately placed at 37 °C. After 30 min, the cytosolic fraction of the cells was prepared, proteins were separated by SDS-PAGE on 10% acrylamide gels under non-reducing conditions and transferred to polyvinylidene difluoride membranes. G-proteins were identified in immunoblot with anti-Gα-common, anti-Gα_i or anti-Gα_q antibodies.

Received 7 February; accepted 2 June 2000.

1. Peach, M. J. Renin-angiotensin system: biochemistry and mechanisms of action. *Physiol. Rev.* **57**, 313–370 (1977).
2. Dudley, D. T. et al. Subclasses of angiotensin II binding sites and their functional significance. *Mol. Pharmacol.* **38**, 370–377 (1990).
3. Parrat, J. R. Protection of the heart by ischemic preconditioning: mechanisms and possibilities for pharmacological exploitation. *Trends Pharmacol. Sci.* **5**, 19–25 (1994).
4. Soubrier, F. et al. Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proc. Natl Acad. Sci. USA* **85**, 9386–9390 (1988).
5. Abdalla, S., Jarnagin, K., Müller-Esterl, W. & Quittner, U. The N-terminal amino group of [Tyr⁸]bradykinin is bound adjacent to analogous amino acids of the human and rat B₂ receptor. *J. Biol. Chem.* **271**, 27382–27387 (1996).
6. Abdalla, S., Zaki, E., Lother, H. & Quittner, U. Involvement of the amino terminus of the B₂ receptor in agonist-induced receptor dimerization. *J. Biol. Chem.* **274**, 26079–26084 (1999).
7. Hein, L. et al. Overexpression of angiotensin AT₁ receptor transgene in the mouse myocardium produces a lethal phenotype associated with myocyte hyperplasia and heart block. *Proc. Natl Acad. Sci. USA* **94**, 6391–6396 (1997).
8. Borkowski, J. A. et al. Targeted disruption of a B₂ bradykinin receptor gene in mice eliminates bradykinin action in smooth muscle and neurons. *J. Biol. Chem.* **270**, 13706–13710 (1995).
9. Zhang, J., Ferguson, S. S. G., Barak, L. S., Menard, L. & Caron, M. G. Dynamin and β-arrestin reveal distinct mechanisms for G protein-coupled receptor internalization. *J. Biol. Chem.* **271**, 118302–118305 (1996).
10. Ransnas, L. A., Svoboda, P., Jasper, J. R. & Insel, P. A. Stimulation of β-adrenergic receptors of S49 lymphoma cells redistributes the α subunit of the stimulatory G protein between cytosol and membranes. *Proc. Natl Acad. Sci. USA* **86**, 7900–7903 (1989).
11. MacEwan, D. J., Kim, G. D. & Milligan, G. Analysis of the role of receptor number in defining the intrinsic activity and potency of partial agonists in neuroblastoma x glioma hybrid NG108-15 cells transfected to express differing levels of the human β₂-adrenoceptor. *Mol. Pharmacol.* **48**, 316–325 (1995).
12. Masaki, H. et al. Cardiac-specific overexpression of angiotensin II AT₂ receptor causes attenuated response to AT₁ receptor-mediated pressor and chronotropic effects. *J. Clin. Invest.* **101**, 527–537 (1998).
13. Nardone, J. & Hogan, P. G. Delineation of a region in the B₂ bradykinin receptor that is essential for high-affinity agonist binding. *Proc. Natl Acad. Sci. USA* **91**, 4417–4421 (1994).
14. Prado, G. N., Taylor, L. & Polgar, P. Effects of intracellular tyrosine residue mutation and carboxyl terminus truncation on signal transduction and internalization of the rat bradykinin B₂ receptor. *J. Biol. Chem.* **272**, 14638–14642 (1997).
15. Hebert, T. E. et al. A peptide derived from a β₂-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J. Biol. Chem.* **271**, 16384–16392 (1996).
16. Jordan, B. A. & Devi, L. A. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature* **399**, 697–700 (1999).
17. Romano, C., Yang, W. -L. & O'Malley, K. L. Metabotropic glutamate receptor 5 is a disulfide-linked dimer. *J. Biol. Chem.* **271**, 28612–28616 (1996).
18. Jones, K. A. et al. GABA_B receptors function as a heteromeric assembly of the subunits GABA_BR1 and GABA_BR2. *Nature* **396**, 674–679 (1998).

19. White, J. H. *et al.* Heterodimerization is required for the formation of a functional GABA_B receptor. *Nature* **396**, 679–682 (1998).
20. Kaupmann, K. *et al.* GABA_B-receptor subtypes assemble into functional heteromeric complexes. *Nature* **396**, 683–687 (1998).
21. Liu, F. *et al.* Direct protein–protein coupling enables cross-talk between dopamine D₅ and γ -aminobutyric acid A receptors. *Nature* **493**, 274–280 (2000).
22. Dixon, B. S., Sharma, R. V., Dickerson, T. & Fortune, J. Bradykinin and angiotensin II: activation of protein kinase C in arterial smooth muscle. *Am. J. Physiol.* **266**, 406–420 (1994).
23. Tsuchida, S. *et al.* Potent antihypertrophic effect of the bradykinin B₂ receptor system on the renal vasculature. *Kidney Int.* **56**, 509–516 (1999).
24. Bascands, J. L. *et al.* Bradykinin-induced *in vitro* contraction of rat mesangial cells via a B₂ receptor type. *Am. J. Physiol.* **267**, F871–F878 (1994).
25. Zhuo, J. *et al.* Localization and interactions of vasoactive peptide receptors in renomedullary interstitial cells of the kidney. *Kidney Int.* **67**, S22–S28. (1998).
26. Quitterer, U. & Lohse, M. J. Crosstalk between G α_i - and G α_q -coupled receptors is mediated by G $\beta\gamma$ exchange. *Proc. Natl Acad. Sci. USA* **96**, 10626–10631 (1999).
27. Quitterer, U., Zaki, E. & Abdalla, S. Investigation of the extracellular accessibility of the connecting loop between membrane domains I and II of the bradykinin B₂ receptor. *J. Biol. Chem.* **274**, 14773–14778 (1999).
28. Mastro, R. & Hall, M. Protein delipidation and precipitation by tri-n-butylphosphate, acetone, and methanol treatment for isoelectric focusing and two-dimensional gel electrophoresis. *Anal. Biochem.* **273**, 313–315 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank B. Nürnberg, for anti-G α -common antibodies, and J. Heukeshoven for helpful advice on high-resolution electrophoresis of hydrophobic proteins. This work was supported in part by the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to U.Q. (e-mail: toph029@rzbox.uni-wuerzburg.de).

Pilus retraction powers bacterial twitching motility

Alexey J. Merz^{*†}, Magdalene So^{*} & Michael P. Sheetz^{†‡}

^{*}Department of Molecular Microbiology and Immunology, Oregon Health Sciences University, Portland, Oregon 97201-3098, USA
[†]Department of Cell Biology, Duke University Medical School, Durham, North Carolina 27705, USA

Twitching and social gliding motility allow many Gram negative bacteria to crawl along surfaces, and are implicated in a wide range of biological functions¹. Type IV pili (Tfp) are required for twitching and social gliding, but the mechanism by which these filaments promote motility has remained enigmatic^{1–4}. Here we use laser tweezers⁵ to show that Tfp forcefully retract. *Neisseria gonorrhoeae* cells that produce Tfp actively crawl on a glass surface and form adherent microcolonies. When laser tweezers are used to place and hold cells near a microcolony, retractile forces pull the cells toward the microcolony. In quantitative experiments, the Tfp of immobilized bacteria bind to latex beads and retract, pulling beads from the tweezers at forces that can exceed 80 pN. Episodes of retraction terminate with release or breakage of the Tfp tether. Both motility and retraction mediated by Tfp occur at about 1 $\mu\text{m s}^{-1}$ and require protein synthesis and function of the PilT protein. Our experiments establish that Tfp filaments retract, generate substantial force and directly mediate cell movement.

Type IV pili are implicated in motility^{1–3,6}, biofilm formation⁷, virulence^{8–11} and all three modes of prokaryotic horizontal genetic transfer (transformation^{12,13}, conjugation¹⁴ and transduction^{15,16}). The Tfp fibre, a helical polymer of the pilin protein, is 6 nm in

diameter and up to several micrometres in length^{17,18}. Type IV pilus biosynthesis occurs through the type II protein translocation pathway and requires several accessory proteins in addition to the pilin subunit¹. The PilT protein is dispensable for Tfp biosynthesis but is essential for both Tfp-mediated motility^{3,19,20} and for the DNA uptake step of genetic transformation^{13,20}. PilT belongs to a highly conserved family of presumed ATPases, which partition to the inner membrane and cytosol and are thought to energize type II and IV protein translocation systems^{1,13,19,21}. Two proteins of this family form hexameric rings strikingly similar to the rings formed by many AAA-type ATP-dependent chaperones and proteases²². Electron-microscope studies of *pilT* mutants led to the hypothesis that Tfp promote cellular motility by retracting^{3,15}, perhaps through PilT-mediated filament disassembly^{1,4,19}. Indirect evidence further suggests that other bacterial surface filaments might retract, including the conjugative F pili of *Escherichia coli*²³ and a type III export filament of *Salmonella enteritidis* associated with cell contact²⁴. However, filament retraction has been neither observed directly nor proven to power motility in any prokaryotic system.

N. gonorrhoeae, the causative agent of gonorrhoea, provides an excellent model for studies of twitching motility. It lacks rotary flagella and components for type III export, and it produces only one known fimbrial structure, the Tfp. When we suspended Tfp-producing *N. gonorrhoeae* cells at low density in liquid medium, they attached to and crawled over the surface of a glass coverslip (Fig. 1). Under optimal conditions (see Methods), over half of the bacteria on the coverslip were motile. Figure 1b depicts the path of a crawling diplococcus (a joined pair of cells is the neisserial functional unit). Although most movements were short and directional changes occurred frequently, many directed movements of 2–5 μm were observed. Cells crawled at $\sim 1 \mu\text{m s}^{-1}$ (Fig. 1c, d). This motility was not due to passive diffusion. Motile cells consistently crawled

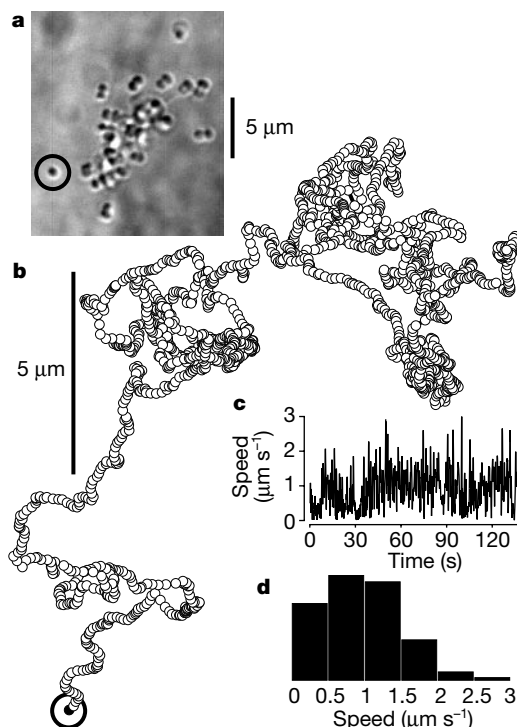


Figure 1 Pilated *N. gonorrhoeae* cells crawl on an inert surface. **a**, Cells crawling on a glass coverslip. This is the first frame of the sequence analysed in **b–d**, and the tracked diplococcus is circled. Note that most cells are present as diplococci. **b**, Tracking of the diplococcus indicated in **a** during an interval of 140 s. Small circles show the position of the tracked diplococcus at intervals of 67 ms. Large circle indicates the start point. **c**, Plot of velocity against time for the same track. **d**, Histogram of the velocities shown in **c**.

[†] Present addresses: Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03755-3844, USA (A.J.M.); Department of Biological Sciences, Columbia University, New York, New York 10027, USA (M.P.S.).

out of a laser-tweezers trap strong enough to firmly arrest suspended cells or 1- μm latex beads (20 pN at 100 nm displacement). Cells crawled at temperatures from 20 to 42 °C. Motility ceased ~ 10 min after the addition of chloramphenicol or tetracycline and resumed upon drug washout, indicating a requirement for protein synthesis (Table 1). Similarly, bacteria were non-motile and unable to grow in defined medium lacking L-glutamine and pyruvate, and became motile upon addition of these nutrients. As expected, non-piliated *pilE* and *pilF* mutants were completely non-motile. Three piliated *pilT* mutants also failed to undergo large movements ($> 1 \mu\text{m}$) and never crawled out of the laser trap (Table 1). *N. gonorrhoeae* cells thus can crawl over surfaces at rates of $\sim 1 \mu\text{m s}^{-1}$ by an active process requiring protein synthesis, Tfp biogenesis and PilT^{3,19,20}.

Tfp-producing *N. gonorrhoeae* cells not only crawl but aggregate into microcolonies that contain twitching or writhing cells. Cells within 1–5 μm of a microcolony often move into the colony, whereas cells within a microcolony move out of the colony^{2,6–8}. To determine whether retractile forces would pull dispersed cells together, we used laser tweezers to position isolated cells 1–5 μm (1–2 pilus lengths) away from microcolonies attached to a coverslip (Fig. 2). The trapped cells were repeatedly pulled from the laser trap toward the microcolonies, directly showing that there are retractile forces between cells (Fig. 2a). The cells were pulled towards the microcolonies at speeds of $\sim 1 \mu\text{m s}^{-1}$, the same rate at which cells crawled on coverslips (Figs 1 and 2a). Cells pulled from the laser trap sometimes bound irreversibly to the microcolonies, but more often were released back into the laser trap (Fig. 2a). Upon release, the tethers between cells were broken, as trapped cells could be moved freely away from the microcolony using the laser tweezers. When released cells were moved $\sim 10 \mu\text{m}$ away and then placed near the opposite side of the same microcolony, they were again pulled out of the trap toward the microcolony (Fig. 2b). Retraction and tethering are therefore transient and can be re-established. In contrast, when we carried out identical manipulations using piliated, non-motile *pilT* mutants, cells were never pulled from the trap (Table 1), although nonretractile static tethers between the trapped cells and the microcolonies sometimes formed (data not shown). Together these experiments show that episodes of PilT-dependent Tfp retraction pull bacterial cells towards one another over distances of up to 5 μm at speeds of $\sim 1 \mu\text{m s}^{-1}$.

Quantitative analyses of Tfp retraction are impeded by variations in cell morphology and by experimental geometries in which two groups of cells pull toward one another. We therefore developed a bead-based assay for Tfp retraction (Fig. 3a). Individual diplococci were immobilized on 3- μm latex beads that had first been coated with anti-*N. gonorrhoeae* antiserum and anchored to the coverslip. Smaller 1- μm beads, coated with a monoclonal antibody that recognizes a surface-exposed epitope on the Tfp fibre, were introduced into the sample chamber. Laser tweezers were then used to hold the 1- μm beads near the immobilized diplococci so that the beads could interact with Tfp.

Table 1 Motility and retraction phenotypes of *N. gonorrhoeae* strains

	Crawling	Cell–cell retraction	Cell–bead retraction
MS11 N400	Yes	Yes	n.d.
MS11 GT102 <i>pilT</i> _{ΔQSL}	No	No	n.d.
MS11 GT103 <i>pilT</i> ::mTnEGNS	No	No	n.d.
MS11 AM92	Yes	Yes	Yes
MS11 AM92T1 <i>pilT</i> ::mTnEGNS	No	No	No
MS11 AM92			
+ Tetracycline	No	n.d.	No
+ Chloramphenicol	No	n.d.	No
–L-glutamine and pyruvate	No	n.d.	n.d.
20 °C (cell-division block)	Yes	n.d.	Yes

n.d., not determined.

When anti-pilus 1- μm beads were placed within 1–3 μm of immobilized cells, the beads became dynamically tethered and were repeatedly pulled towards the immobilized cells (Fig. 3b). The mean speed of retraction was $1.17 \pm 0.49 \mu\text{m s}^{-1}$ (mean $\pm$ s.d.; $n = 713$), and was independent of the distance traversed by the beads (Fig. 3c). Retraction events were separated by intervals of 1–20 s. One-micrometre beads lacking the anti-pilin monoclonal antibody did not bind to pili as frequently as anti-pilin beads; but when they bound, they also were pulled toward immobilized bacteria. Type IV pili therefore exert retractile force through both nonspecific and specific binding interactions, consistent with our finding that Tfp facilitate bacterial motility on inert substrates (Fig. 1). Retraction ceased ~ 10 min after the addition of chloramphenicol or tetracycline and resumed after drug washout, again indicating a requirement for protein synthesis.

The restoring force of the laser tweezers increases with radial displacement from the trap centre, and can be calibrated for homogeneous particles using laminar flow (see Methods). Figure 3c shows a trace of displacement against time, with force shown on the displacement axis. Isolated, immobilized diplococci were often able to exert forces of 80 pN or more on trapped 1- μm beads. By comparison, ~ 20 pN is the force needed to extract an integral membrane protein from a lipid bilayer²⁵, and ~ 30 pN of tensile force is sufficient to elongate a microvillus on a host cell²⁶.

Retraction usually terminated with release or breakage of the pilus tether and bead movement back into the laser trap (Fig. 3b, c). Transitions from retraction to release usually occurred in less than 67 ms. The velocities of bead movement back into the trap (Fig. 3c) were much more variable than retraction speeds and were proportional to the distance displaced before release ($4.34 \pm 3.48 \mu\text{m s}^{-1}$;

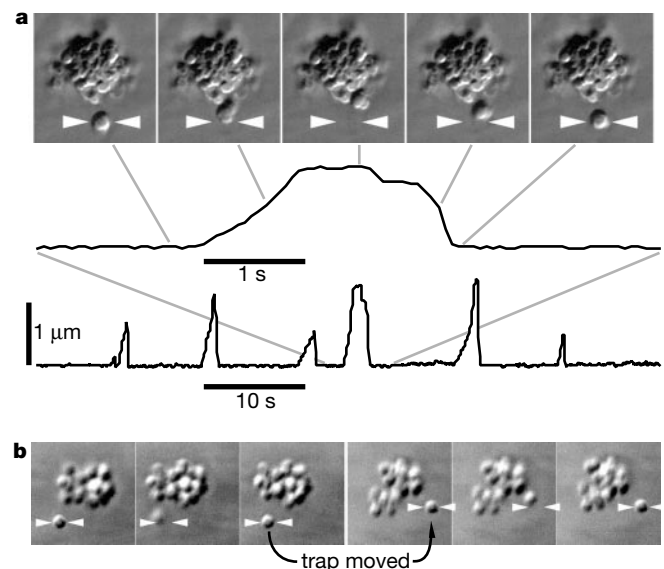


Figure 2 Optical tweezers reveal retractile forces between piliated *N. gonorrhoeae* cells. **a**, Cell–microcolony assay. A diplococcus held in the laser-tweezers trap (indicated by arrowheads) was repeatedly pulled toward a microcolony attached to the coverslip. Micrographs (separated by ~ 1 -s intervals) show one retraction event from the sequence used to generate the trace. Traces indicate radial displacement of the trapped diplococcus from the centre of the laser trap. The top trace shows a time-magnified portion of the bottom trace. Note that retraction occurs at $\sim 1 \mu\text{m s}^{-1}$. **b**, Retraction and tethering are transient and can be re-established. A diplococcus was trapped and positioned next to a microcolony as in **a**. The diplococcus was pulled from the trap towards the microcolony and then released back into the trap (first three panels). Next, the trap and trapped diplococcus were moved more than 10 μm away from the microcolony, then repositioned on the opposite side of the same microcolony, where the diplococcus was again pulled from the trap towards the microcolony (last three panels).

$n = 171$). Release events could result from force-regulated release of the fibre at a threshold tension level, from breakage of the retracting fibre or the basal body, or from dissociation of the retracting fibre from the 1- μm bead. More detailed analyses of many traces made at laser power levels of 400 and 800 mW (Fig. 3e) indicated that bead release is dependent on force, and is not solely a function of linear displacement. Histograms of the forces reached before bead release show small peaks at about 40, 60 and 90 pN (Fig. 3e, arrows). Studies of other motor-dependent movements have also revealed peaks in displacement histograms. These peaks have been interpreted as multiples of the critical force generated by the motors²⁷. The peaks observed in our study may reflect the presence of multiple retracting fibres or multiple motor units on a single fibre; however, we stress that additional, more refined measurements will be required to establish the unit critical force of the Tfp retraction machinery.

As in the cell–microcolony assays, *pilT* mutant cells were unable to generate retractile forces but could sometimes form static tethers to laser-trapped 1- μm beads. When a tethered bead was held in a stationary laser-tweezers trap, movement of the microscope stage

(and hence the immobilized bacterium) within a defined radius did not displace the tethered 1- μm bead out of the trap, but movement outside this radius displaced the tethered bead from the trap centre (Fig. 3f). The static character of the tether was tested by placing the tethered bead under tension, 0.2 μm from the centre of the trap. No further displacements of the bead (> 25 nm) occurred over a period of more than 120 s, indicating that the tether was static (Fig. 3f). These results confirm that PilT is essential for Tfp retraction.

What is the mechanism of Tfp retraction? The available evidence is most consistent with molecular ratchet models²⁸ in which retractile force is generated by filament disassembly into the inner membrane. Genetic analyses suggest that PilT promotes pilus disassembly, pilin degradation, or both^{4,19}. PilT, like the closely related ATPases thought to energize Tfp assembly and type II and IV secretion, localizes to the cytoplasm and inner membrane^{1,13,19,21}. Unassembled pilin is a type II integral inner membrane protein²⁹. Upon assembly, the membrane-spanning domain of pilin moves into a hydrophobic coiled-coil at the core of the polymeric fibre^{17,18}. These observations support models in which cytoplasmic ATPases

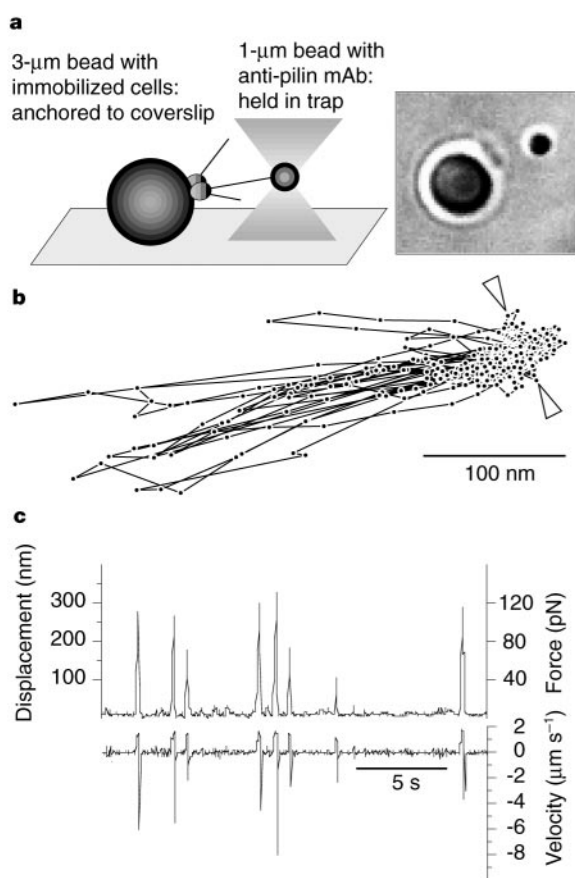
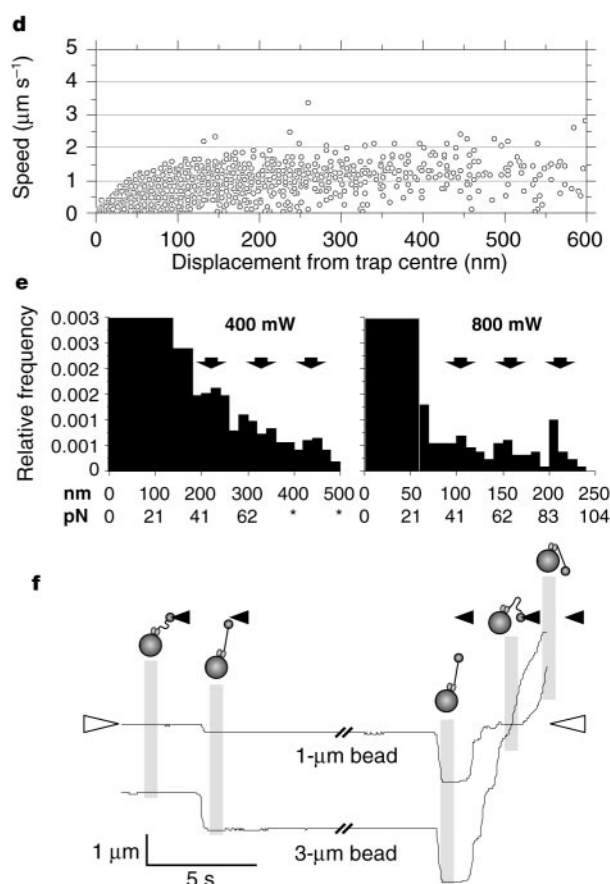


Figure 3 Quantitative type IV pili retraction assay. **a**, Experimental geometry. The cartoon shows a side view; the micrograph shows the assay in progress from above, with scale indicated by the beads (1 and 3 μm). **b**, Direction of displacements. Trace from a recording of a single immobilized diplococcus that shows the locations of the 1- μm bead centroid in the plane parallel to the coverslip surface. The centre of the laser trap is indicated (arrowheads). The immobilized diplococcus is in the same position relative to the laser trap as in the micrograph in **a**. **c**, Velocity, timing and force of retraction. Top trace, radial displacement of the 1- μm bead versus time, and calibrated restoring force imposed by the laser-tweezers trap. Bottom trace, velocities of retraction and release from the same recording. Positive velocities denote displacement away from the trap centre (toward the immobilized cells), and negative velocities denote release into the trap centre. **d**, Plot of retraction velocity versus displacement, showing pooled data from recordings of



14 different immobilized diplococci. Note that for displacements of 100 nm or more, retraction velocities are relatively constant at $\sim 1 \mu\text{m s}^{-1}$. **e**, Histograms of forces reached prior to bead release at 400 and 800 mW input laser power. **f**, Static tethering of 1- μm beads by immobilized *pilT* mutant cells. Cartoons show interpretation of the traces. The Tfp could not be seen, but their state (slack or under tension) was inferred from movements of the 1- μm bead relative to the trap. Arrowheads denote the position of the laser trap (filled for the cartoons, open for the trace), which was stationary throughout the experiment. The 3- μm bead was attached to the coverslip on the microscope stage, and was moved using a stage motor. Traces show the displacements of the 1- μm and 3- μm beads relative to the laser trap plotted against time. Breaks in the traces denote an interval of more than 120 s in which there were no movements more than 25 nm of the 1- μm bead.

catalyse pilus extension and retraction from the base of the pilus fibre, within the plane of the inner membrane. The Tfp filament is a single-stranded helix with five pilin subunits per turn and a 40 Å pitch^{17,18}. Assembly and disassembly are therefore predicted to occur in 8 Å steps, consistent with the high levels of retractile force observed (Figs. 3c, e). Retraction mediated by disassembly at the observed rate of 1.2 µm s⁻¹ (Fig. 3d) would entail the removal of 1,500 pilin subunits per second from the fibre base.

Pilus retraction is thought to occur in several bacterial systems^{3,4,15,19,23,24}. To our knowledge our experiments provide the first direct observations of pilus retraction and may provide a model for understanding related retractile processes. In *N. gonorrhoeae*, DNA transformation facilitates pilin antigenic variation¹² and other horizontal genetic transfer functions. Like twitching motility, DNA uptake requires Tfp biosynthesis and PilT function^{13,20}. Moreover, many bacteria use PilT or PilT orthologues for DNA uptake, including species that lack Tfp. Consistent with the hypothesis that DNA uptake and pilus retraction use a common machinery, the rate of processive DNA uptake in *Haemophilus influenzae* (~0.17 µm s⁻¹; ref. 13) is comparable to the speed of Tfp retraction in our experiments. A broadly conserved mechanism may therefore power not only pilus retraction, but also other macromolecular translocation processes. Finally, PilT and twitching motility are implicated in virulence-related functions^{8–10}. Eukaryotic cells sense and respond to mechanical forces³⁰, and *N. gonorrhoeae* Tfp and PilT function together to elicit adhesive plaque formation in epithelial cells¹¹; we therefore propose that the mechanical forces generated by retractile pili are important and previously overlooked signals between pathogenic bacteria and host cells. □

Methods

Bacterial strains and media

N. gonorrhoeae were maintained on GCB agar with supplements (Difco). Antibiotics (Sigma) were used at the following concentrations (in mg l⁻¹): kanamycin, 100; erythromycin, 4; tetracycline, 50; chloramphenicol, 40. We carried out motility assays in phenol red-free DMEM (Gibco BRL), supplemented with L-glutamine, pyruvate and 1% (w/v) bovine serum albumin (BSA) fraction V (Sigma). BSA limited the adsorption of bacteria onto the glass coverslips and was essential for optimal motility. MS11 N400 and its isogenic derivatives GT102 (*pilTΔQSL*) and GT103 (*pilT::mTnEGNS*) were gifts from M. Koomey²⁰. MS11 AM92 is an MS11A derivative that produces a pilin variant recognized by monoclonal antibody 20D910 (ref. 10). The *pilT::mTnEGNS* allele from GT103 was moved into MS11 AM92 by DNA transformation, then backcrossed against MS11 AM92 three times to generate AM92T1. AM92 and AM92T1 are pilated as judged by immunofluorescent staining with monoclonal antibody 20D9 and by immunoblotting with monoclonal antibodies 20D9 and 10H5 (anti-SM1 epitope). As expected, AM92 is pilated, motile, competent for DNA transformation, and elicits cortical plaques in epithelial cells, whereas AM92T1 is pilated, non-motile, transformation deficient and exhibits defective cortical plaque induction^{11,12,20}.

Bead-based retraction assay

Three-micrometre latex beads (Bangs Laboratories) were coated with rabbit anti-*N. gonorrhoeae* antiserum by passive adsorption as recommended by the supplier. The beads were washed thoroughly in PBS (phosphate-buffered saline, pH 7.4) and adsorbed onto clean coverslips in the presence of PBS. Coverslips were washed in PBS to remove unbound beads and blocked in DMEM with 1% BSA at 25 °C for 15 min or more before use. One-micrometre beads were coated with biotinylated monoclonal antibody 20D9 by a sandwich method³⁰. One-micrometre beads (Polysciences) were covalently coupled to biotinylated casein (Sigma). Biotinylated monoclonal antibody 20D9 (prepared with sulpho-NHS biotin; Pierce) was then linked to the biotinylated casein by an avidin bridge; control beads lacked monoclonal antibody 20D9. The beads were blocked with biotinylated BSA (biotinamido caproyl BSA; Sigma) and washed thoroughly. A dot blot assay indicated that roughly 5 × 10⁴ molecules of monoclonal antibody 20D9 bound per bead. A serial perfusion method greatly reduced aggregate formation among the beads and cells. Coverslips with attached 3-µm beads were used to form the bottom surface of a perfusion chamber, which was placed on the microscope stage. A dilute, vortex-dispersed suspension of *N. gonorrhoeae* cells in DMEM with 1% BSA was perfused into the chamber. We monitored bacterial binding to the anchored 3-µm beads by microscope. When 10–20% of the 3-µm beads had bound 1 or 2 diplococci, unbound bacteria were washed out. One-micrometre beads were suspended in DMEM with 1% BSA and dispersed by bath sonication, then perfused into the sample chamber and manipulated using the laser tweezers. Except where indicated, we carried out assays at 35 °C.

Laser trapping and single particle tracking

The optical tweezers workstation was configured as described⁵ and included a Zeiss Axiocvert microscope, a Nd-YAG laser (1,064-nm wavelength) directed into the bottom port of the microscope, a piezoelectric stage controller, and a Nuicon camera with analogue signal processor and S-VHS video recorder for data capture. We carried out video digitization and data analysis on a SGI workstation running Isee particle tracking software (Inovision). Speeds of crawling bacteria (Fig. 1c, d) were obtained by calculating centroid position using a moving average function with a three-frame window to minimize artefacts arising from the irregular shapes of the cells. In the bead-based assay, velocities were calculated frame by frame. Positive velocities were assigned to displacements away from the trap centre (retraction) and negative velocities were assigned to movement toward the trap centre (release). Values less than three times the r.m.s. retraction or release velocities were discarded as noise. The remaining values were used to calculate the mean retraction and release velocities shown in the text. Results obtained by this method agreed with results obtained by hand-fitting the slopes of displacement versus time plots for individual retraction events. The trap force calibration was done as described⁵.

Received 28 April; accepted 20 June 2000.

- Wall, D. & Kaiser, D. Type IV pili and cell motility. *Mol. Microbiol.* **32**, 1–10 (1999).
- Henrichsen, J. Twitching motility. *Annu. Rev. Microbiol.* **37**, 81–93 (1983).
- Bradley, D. E. A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. *Can. J. Microbiol.* **26**, 146–154 (1980).
- Wolfgang, M., Park, H. S., Hayes, S. F., van Putten, J. P. M. & Koomey, M. Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function mutations in *pilT*, a twitching motility gene in *Neisseria gonorrhoeae*. *Proc. Natl Acad. Sci. USA* **95**, 14973–14978 (1998).
- Sheetz, M. P. (ed.) *Laser Tweezers in Cell Biology* (Academic, New York, 1997).
- Swanson, J. Studies on gonococcus infection. XII. Colony color and opacity variants of gonococci. *Infect. Immun.* **19**, 320–331 (1978).
- O'Toole, G. A. & Kolter, R. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* **30**, 295–304 (1998).
- Bieber, D. et al. Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* **280**, 2114–2118 (1998).
- Comolli, J. C. et al. *Pseudomonas aeruginosa* gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. *Infect. Immun.* **67**, 3625–3630 (1999).
- Pujol, C., Eugene, E., Marceau, M. & Nassif, X. The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion. *Proc. Natl Acad. Sci. USA* **96**, 4017–4022 (1999).
- Merz, A. J., Enns, C. A. & So, M. Type IV pili of pathogenic *Neisseriae* elicit cortical plaque formation in epithelial cells. *Mol. Microbiol.* **32**, 1316–1332 (1999).
- Seifert, H. S., Ajioka, R. S., Marchal, C., Sparling, P. F. & So, M. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. *Nature* **336**, 392–395 (1988).
- Dubnau, D. DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53**, 217–244 (1999).
- Yoshida, T., Kim, S. R. & Komano, T. Twelve *pil* genes are required for biogenesis of the R64 thin pilus. *J. Bacteriol.* **181**, 2038–2043 (1999).
- Bradley, D. E. Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochem. Biophys. Res. Commun.* **47**, 142–149 (1972).
- Karaolis, D. K., Somara, S., Maneval, D. R. Jr, Johnson, J. A. & Kaper, J. B. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* **399**, 375–379 (1999).
- Parge, H. E. et al. Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature* **378**, 32–38 (1995).
- Forest, K. T. & Tainer, J. A. Type-4 pilus structure: outside to inside and top to bottom—a minireview. *Gene* **192**, 165–169 (1997).
- Whitchurch, C. B., Hobbs, M., Livingston, S. P., Krishnapillai, V. & Mattick, J. S. Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. *Gene* **101**, 33–44 (1991).
- Wolfgang, M. et al. *pilT* mutations lead to simultaneous defects in competence for natural transformation and twitching motility in pilated *Neisseria gonorrhoeae*. *Mol. Microbiol.* **29**, 321–330 (1998).
- Brossay, L., Paradis, G., Fox, R., Koomey, M. & Hebert, J. Identification, localization, and distribution of the PilT protein in *Neisseria gonorrhoeae*. *Infect. Immun.* **62**, 2302–2308 (1994).
- Krause, S. et al. Sequence-related protein export NTPases encoded by the conjugative transfer region of RP4 and by the *cag* pathogenicity island of *Helicobacter pylori* share similar hexameric ring structures. *Proc. Natl Acad. Sci. USA* **97**, 3067–3072 (2000).
- Novotny, C. P. & Fives-Taylor, P. Retraction of F pili. *J. Bacteriol.* **117**, 1306–1311 (1974).
- Giococchio, C. C., Olmsted, S. B., Wells, C. L. & Galan, J. E. Contact with epithelial cells induces the formation of surface appendages on *Salmonella typhimurium*. *Cell* **76**, 717–724 (1994).
- Evans, E., Berk, D. & Leung, A. Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments. *Biophys. J.* **59**, 838–848 (1991).
- Shao, J. Y., Ting-Beall, H. P. & Hochmuth, R. M. Static and dynamic lengths of neutrophil microvilli. *Proc. Natl Acad. Sci. USA* **95**, 6797–6802 (1998).
- Coppin, C. M., Finer, J. T., Spudich, J. A. & Vale, R. D. Detection of sub-8-nm movements of kinesin by high-resolution optical-trap microscopy. *Proc. Natl Acad. Sci. USA* **93**, 1913–1917 (1996).
- Mahadevan, L. & Matsudaira, P. Motility powered by supramolecular springs and ratchets. *Science* **288**, 95–100 (2000).
- Dupuy, B., Taha, M. K., Pugsley, A. P. & Marchal, C. *Neisseria gonorrhoeae* prepilin export studied in *Escherichia coli*. *J. Bacteriol.* **173**, 7589–7598 (1991).
- Felsenfeld, D. P., Schwartzberg, P. L., Venegas, A., Tse, R. & Sheetz, M. P. Selective regulation of integrin-cytoskeleton interactions by the tyrosine kinase Src. *Nature Cell Biol.* **1**, 200–206 (1999).

Acknowledgements

We thank our colleagues in the Sheetz and So labs for invaluable technical assistance and

stimulating discussions; E. Barklis and L. Kenney for critical comments on the manuscript; and M. Koomey for providing bacterial strains. This work was supported by NIH grants to M.S. and M.P.S. A.J.M. received pre-doctoral support from an NIH NRSA grant and postdoctoral support from the Cancer Research Fund of the Damon Runyan-Walter Winchell Foundation.

Correspondence and requests for materials should be addressed to M.P.S. (e-mail: ms2001@columbia.edu).

TFIIH is negatively regulated by cdk8-containing mediator complexes

Sasha Akoulitchev, Sergei Chuikov & Danny Reinberg*

Howard Hughes Medical Institute, Division of Nucleic Acids Enzymology, Department of Biochemistry, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA

The mammalian cyclin-dependent kinase 8 (cdk8)¹ gene has been linked with a subset of acute lymphoblastic leukaemias², and its corresponding protein has been functionally implicated in regulation of transcription^{3,4}. Mammalian cdk8 and cyclin C, and their respective yeast homologues, Srb10 and Srb11, are components of the RNA polymerase II holoenzyme complex^{5,6} where they function as a protein kinase that phosphorylates the carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II (ref. 7). The yeast *SRB10* and *SRB11* genes have been implicated in the negative regulation of transcription⁸. The cdk8/cyclin C protein complex is also found in a number of mammalian Mediator-like protein complexes^{3,5,9–12}, which repress activated transcription independently of the CTD *in vitro*^{9,10}. Here we show that cdk8/cyclin C can regulate transcription by targeting the cdk7/cyclin H subunits of the general transcription initiation factor IIH (TFIIH). cdk8 phosphorylates mammalian cyclin H in the vicinity of its functionally unique amino-terminal and carboxy-terminal α -helical domains¹³. This phosphorylation represses both the ability of TFIIH to activate transcription and its CTD kinase activity. In addition, mimicking cdk8 phosphorylation of cyclin H *in vivo* has a dominant-negative effect on cell growth. Our results link the Mediator complex and the basal transcription machinery by a regulatory pathway involving two cyclin-dependent kinases. This pathway appears to be unique to higher organisms.

The stimulation of gene-specific transcription by transcriptional activator proteins requires coactivators—molecules that mediate communication between the general transcription machinery and the activators. Some transcription coactivators are activator- or gene-specific, whereas others are required for more global gene transcription (general coactivators). General coactivators include PC4, TFIIA and components of the TFIID and RNA polymerase II (RNAPII) holoenzyme complexes. Coactivators in the RNAPII holoenzyme appear to be modular in nature, forming distinct subcomplexes that are capable of positively or negatively regulating transcription¹². Distinct coactivator complexes have been isolated from mammalian cells. One such complex is NAT⁹, which appears to be functionally similar to the mammalian SMCC–TRAP complex¹⁰ and which contains subsets of polypeptides present in other mammalian coactivator complexes such as Mediator¹¹, DRIP¹⁴, CRSP¹⁵ and ARC¹⁶. A unique feature of NAT, Mediator and SMCC–TRAP complexes is the presence of cdk8 and cyclin C, which are the human homologues of yeast Srb10 and Srb11, respectively. The Srb10/11 complex and the NAT complex down-regulate transcription by phosphorylating the CTD of RNAPII before its association with transcription initiation complexes^{7,9};

however, NAT also downregulates transcription independently of the CTD of RNAPII⁹. Notably, other cdk8-containing complexes, such as SMCC–TRAP and Mediator, can function as coactivators of transcription when the reconstituted system is supplemented with a crude protein fraction¹¹ or when the system is reconstituted with PC4 in the absence (or with limiting amounts) of TFIID¹⁰.

To investigate further the mechanism by which cdk8 negatively regulates transcription, we isolated the NAT and SMCC–TRAP complexes^{9,10}, and analysed their effects on transcription as a function of TFIID (Fig. 1a). In the absence of TFIID, both complexes enhanced VP16-mediated activation of transcription. However, both complexes mediated the repression of activated transcription in the TFIID-dependent assay. Notably, a NAT complex containing a kinase-deficient mutant of cdk8 (D173A) did not support repression (Fig. 1b–d). Treatment of immunoadsorbed, highly purified TFIID with recombinant cdk8/cyclin C under kinase conditions impaired transcription in a phosphatase-reversible manner (Fig. 1e), which indicates that the cdk8 kinase can downregulate transcription by phosphorylating TFIID.

The transcription system described above requires PC4, which is downregulated by phosphorylation¹⁷, and it has been suggested that cdk8 targets PC4 (ref. 10). To deduce the *in vivo* target of cdk8, we reconstituted activated transcription in the absence of PC4 using a coactivator complex devoid of cdk8; cdk8 was added exogenously to the assays. We purified and analysed Mediator-like complexes that were devoid of cdk8 but contained polypeptides found in other coactivator complexes (see Supplementary Information). The Mediator-like complex displayed properties similar to those

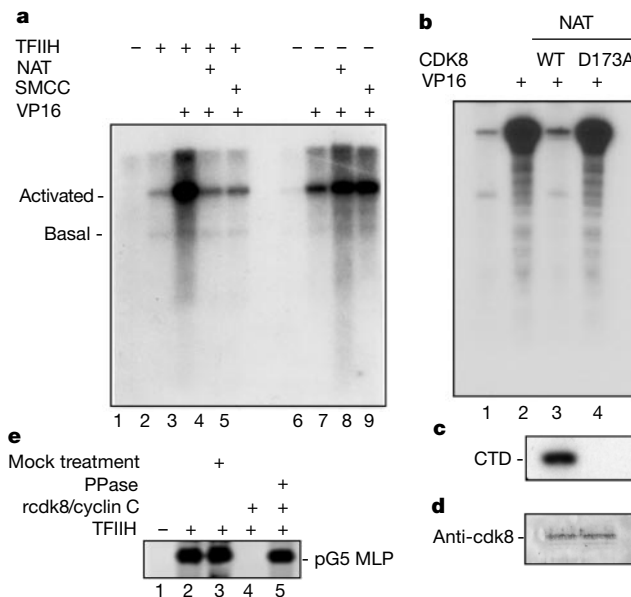
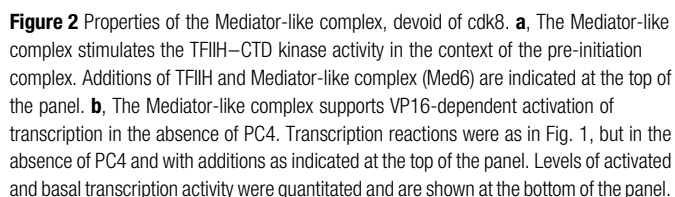
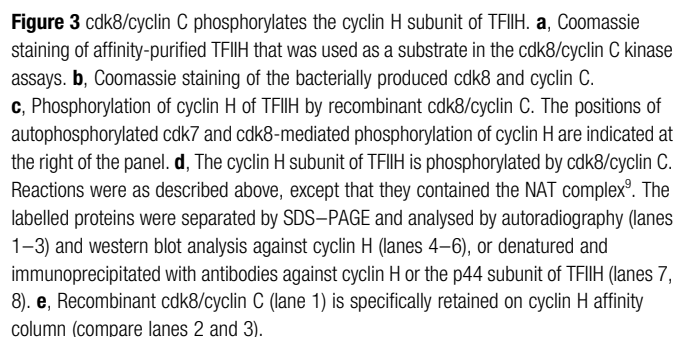


Figure 1 TFIID mediates cdk8-dependent repression by NAT. **a**, TFIID-dependent activities of NAT and SMCC complexes. NAT and SMCC–TRAP complexes were purified as described^{9,10}, and their activities analysed using a reconstituted VP16-dependent transcription activation assay in the presence of PC4. Linear (lanes 1–5) or supercoiled (lanes 6–9) DNA templates were used in the assays, and transcription activity was measured in the presence and absence of TFIID, respectively. The supercoiled template was used to render the assay independent of TFIID²⁹. The positions of the activated and basal transcription products are indicated. **b–d**, Repression by NAT is cdk8 dependent. NAT was immunopurified from 293T cells expressing wild-type cdk8 (WT) or the kinase-deficient mutant³ of cdk8 (D173A). The preparation was assayed by TFIID-dependent activated transcription (**b**), CTD kinase assay (**c**) and western blot (**d**). **e**, Cdk8 phosphorylates and inactivates TFIID. PPase, alkaline phosphatase.

We identified the amino acids of cyclin H that were phosphorylated by recombinant cdk8/cyclin C by using a set of deletions and point mutations of cyclin H spanning all putative phospho-acceptor sites recognized by the CDK family of protein kinases¹⁹ (Fig. 4a; and data not shown). cdk8/cyclin C phosphorylates cyclin H at Ser5 and Ser304 both *in vitro* and *in vivo* (data not shown). These two phosphorylation sites are in the vicinity of N- and C-terminal helices which are uniquely orientated in cyclin H as compared with other cyclins (Fig. 4b). Their disruption results in a loss of cdk7 kinase activity *in vitro*¹³. We reasoned that phosphorylation of Ser5 and/or Ser304 might interfere with the biological activities of TFIIF. Purified TFIIF immuno-adsorbed to beads was treated with cdk8/cyclin C under kinase reaction conditions and then analysed in a transcription assay (Fig. 4c). The cdk8 treatment of TFIIF resulted in inhibition of its transcription activity



We next analysed the effect of mimicking phosphorylation of cyclin H at Ser 5 and/or Ser 304 by aspartic acid substitution. We used an inducible expression system that encodes versions of cyclin H containing a histidine tag at the C terminus (Fig. 5a). The expression vector was maintained episomally in 293EBNA cells under selection with hygromycin B (ref. 20). Expression of Histagged cyclin H was induced with Cd^{2+} from the human metallothionein II promoter²¹. We purified the mutated forms of TFIIH from transfected cells using a double immunopurification procedure²², and we analysed equivalent amounts of the mutated TFIIH proteins (Fig. 5b) for CTD kinase and transcription activity. The TFIIH CTD kinase activity was impaired in the S5D (DS), but not in the S304D (SD) mutation, and the double mutant was impaired as expected (Fig. 5c). The transcription activity of TFIIH was affected by each of the amino-acid substitutions; however, the single substitutions supported transcription, whereas the double mutant displayed the most severe impairment (Fig. 5d). The regulation of TFIIH transcription activity by cdk-activating kinase (CAK) is in agreement with previous studies indicating that the



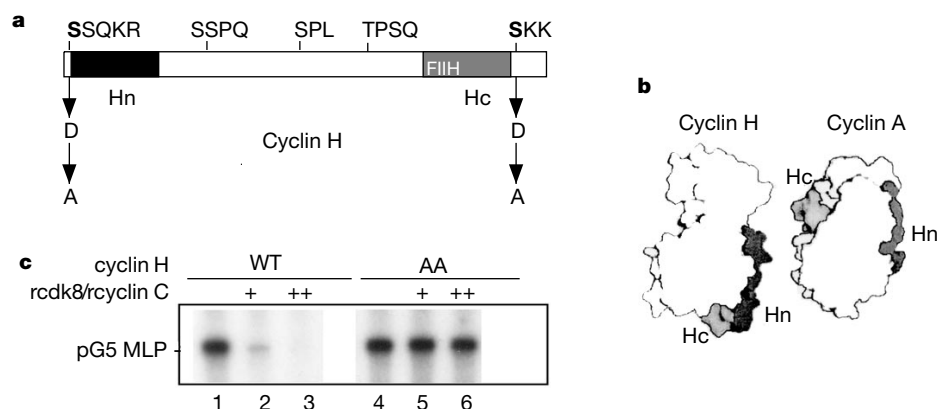


Figure 4 Phosphorylation of cyclin H by cdk8 at Ser 5 and Ser 304 *in vitro*.

a, Representation of cyclin H and putative phosphorylation sites analysed *in vitro*.

b, Molecular surface of cyclin H and cyclin A (ref. 13). Hn and Hc of mammalian cyclin H correspond to the N- and C-terminal α -helices. Notice that Hn and Hc are brought

together in cyclin H, a feature unique to mammalian cyclin H¹³. **c**, cdk8-mediated phosphorylation of cyclin H inactivates TFIIH. Immunopurified TFIIH containing wild-type (lanes 1–3) or double alanine mutant (AA) cyclin H (lanes 4–6) was treated with cdk8/cyclin C and analysed directly in transcription reactions.

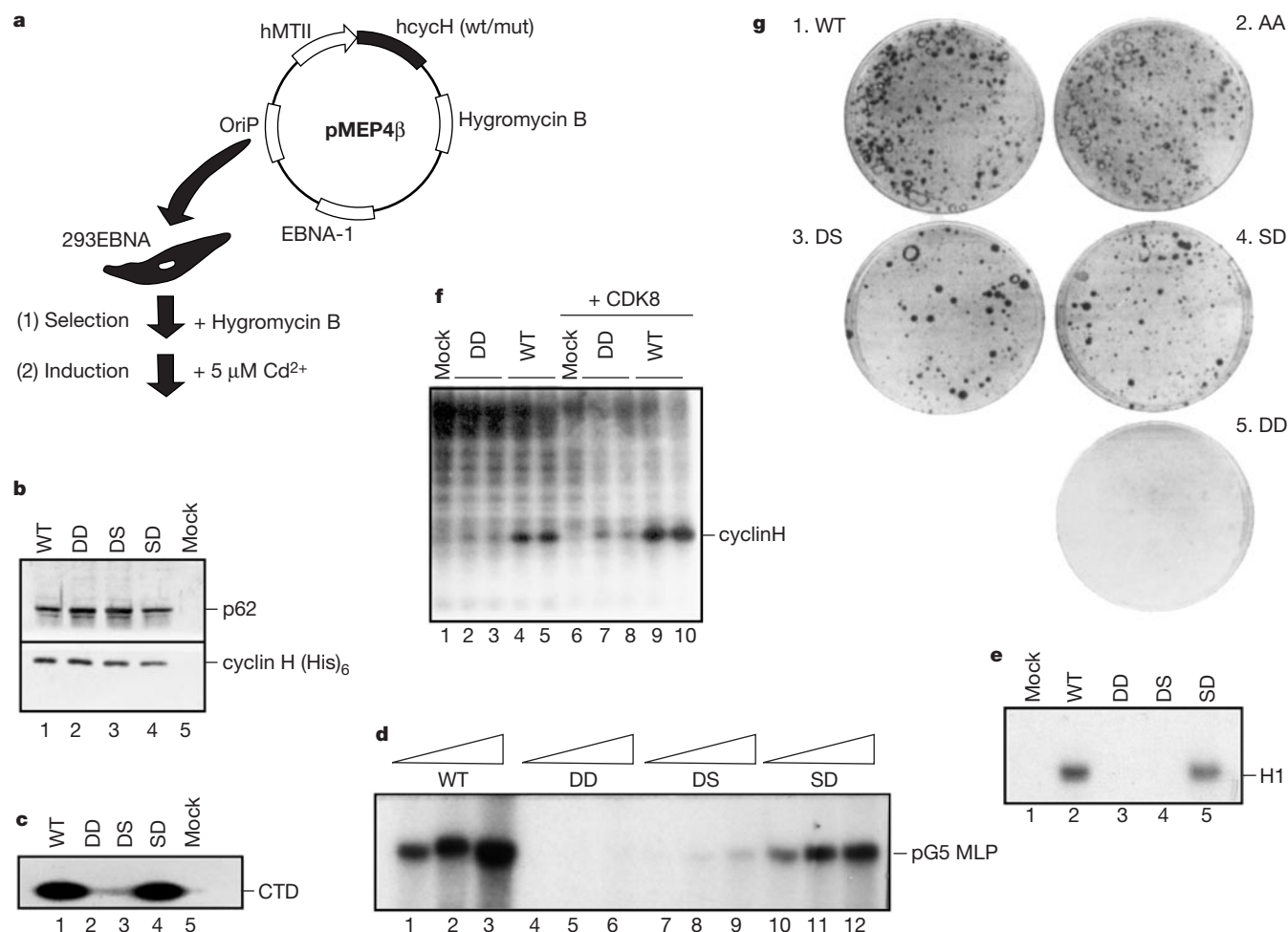


Figure 5 Regulatory phosphorylation of cyclin H *in vivo*. **a**, Inducible system used for the expression of wild-type and mutant cyclin H (ref. 20). **b**, Expression and purification of TFIIH with mutant forms of cyclin H. Western blot of immunopurified TFIIH containing wild-type cyclin H (WT), the double mutant S5D,S304D (DD), the S5D mutant (DS) or the S304D mutant (SD), is shown. The p62 and cyclin H subunits of TFIIH are indicated. **c**, CTD kinase assays with mutants of TFIIH and with a CTD peptide used as substrate. **d**, Reconstituted transcription assay with the mutants of TFIIH. The migration position of the specific transcript is indicated (pG5 MLP). **e**, CAK assay with the mutants of TFIIH. **f**, Ser 5 and

Ser 304 of cyclin H are phosphorylated *in vivo* in a manner dependent on cdk8. His-tagged cyclin H and its double mutant (DD) were labelled *in vivo* upon induction (lanes 2–10) with or without transiently coexpressed cdk8/cyclin C (lanes 7–10). The experiment was carried out in duplicate and normalized for protein and cyclin H content (not shown). The migration position of immunopurified tagged cyclin H on the autoradiogram is indicated. **g**, Cellular growth is inhibited upon induction of cyclin H mutants. Plates with 293EBNA cells carrying the indicated cyclin H expression vectors were stained after two weeks of growth under inducible conditions.

CAK complex regulated TFIIF transcription activity through an allosteric mechanism^{22,23}. We also analysed the CAK complex containing cyclin H mutants (purified from the transfected cells) for their ability to induce activation of cdc2 (Fig. 5e)²⁴. Like CTD-kinase activity, the double mutant and the S5D were impaired in activation of cdc2-cyclin B, as measured using histone H1 as a substrate.

In vivo labelling of the transfected 293EBNA cells with ³²P-orthophosphate upon induction of cyclin H expression showed that cyclin H is phosphorylated *in vivo* (Fig. 5f, lanes 4, 5). Moreover, its phosphorylation levels were induced upon transient expression of co-transfected cdk8/cyclin C (Fig. 5f, compare lanes 4, 5 and 9, 10). Phosphorylation of the double aspartate mutant of cyclin H was reduced markedly and did not change upon coexpression of cdk8/cyclin C. Together with peptide mapping data and western blot controls (data not shown), these results indicate that Ser 5 and Ser 304 are phosphorylated *in vivo* in response to cdk8/cyclin C.

The mammalian expression system allowed us to analyse the effect of cyclin H substitutions on cellular proliferation (Fig. 5g). We plated the hygromycin B-selected pools of 293EBNA cells carrying the mutant forms of cyclin H at low density and allowed them to grow for two weeks in 200 nM Cd²⁺; the cells were then fixed and stained. Sub-optimal concentrations of Cd²⁺ reduced the long-term toxic effect of Cd²⁺. The cyclin H double mutant showed a dominant-negative effect on cellular growth, with no colonies formed after 14 days of incubation. This effect was specific for the amino-acid substitutions that mimicked phosphorylation, as substitution of the same residues with alanine did not inhibit cellular growth. Together, our results establish that cdk8/cyclin C induces phosphorylation of cyclin H *in vivo* at serine residues 5 and 304, resulting in the downregulation of TFIIF activities and, consequently, in the inhibition of cellular proliferation. The growth inhibition of the mutants correlates with the transcription data shown above. Single substitutions appear to have a less severe effect, whereas the double substitution has a severe effect on transcription and failed to support cell growth. However, the results do not rule out that the effect observed on cellular growth is solely due to TFIIF transcription activity, and not due to the CAK complex affecting the cell cycle.

This cdk8/cyclin C/cyclin H regulatory pathway appears to be unique to metazoans. We have not been able to demonstrate Srb10/11-mediated phosphorylation of Ccl1 in *Saccharomyces cerevisiae* (also noted by R. Young, personal communication). Although most of the essential processes between yeast and higher eukaryotes are well conserved, it is not surprising that differences between unicellular and multicellular organisms exist in the regulation of these essential processes. In higher eukaryotes, cdk8/cyclin C may have evolved to acquire a more intricate regulatory function in response to the increased requirement for expression and regulation of more complex genomes. In fact, cdk7/cyclin H, the target of cdk8/cyclin C, features additional regulatory function in higher eukaryotes as a CAK²⁴. The yeast homologue Kin28/Ccl1 does not support CAK activity.

The negative regulatory functions of cdk8/cyclin C that operate in higher eukaryotes exert their effects at two crucial steps in the transcription-initiation pathway. By phosphorylating the CTD early in the transcription-initiation process before RNAPII has been assembled into a transcription-initiation complex, cdk8/cyclin C can halt the formation of such a complex. The second target of cdk8/cyclin C, TFIIF, is required for completion of the transcription-initiation process, namely, open complex formation and promoter escape²⁵. Thus, if RNAPII has associated with the transcription-initiation complex, and the transcription requirements of various genes change according to environmental conditions, then cdk8/cyclin C can inhibit transcription by regulating the activities of TFIIF. □

Methods

Transcription

Transcription was done as described²². The factors used in the reconstituted system were recombinant bacterially expressed Gal4–VP16, PC4, TFIIB, TFIIF, TFIIE; affinity-purified TFIIF and TFIID; and highly purified mammalian RNAPII. Reactions contained two different DNA templates that differed in the length of the G-less cassette and in the presence of Gal4-binding sites upstream of the TATA motif of the adenovirus major late promoter. The DNA with the shorter cassette was devoid of Gal4-binding sites and measured basal transcription.

Purification of the NAT complex containing tagged cdk8

His-tagged wild-type or mutant cdk8 was transiently expressed from the pcDNA3.1 plasmid. We purified NAT complex by conventional chromatography⁹ followed by immunopurification from the DE-52 flow through fraction with anti-His-tag monoclonal antibodies.

Inactivation of TFIIF by cdk8/cyclin C

TFIIF was immobilized on beads using monoclonal antibodies against ERCC3. The immuno-adsorbed TFIIF was treated with recombinant cdk8/cyclin C in the presence of ATP under kinase reaction conditions²². After treatment, TFIIF was washed to remove cdk8/cyclin C, and a fraction of the beads were analysed directly in transcription reactions. We treated an equivalent fraction with alkaline phosphatase as described²². After a wash (0.5M NaCl, 0.1% NP40) to remove the phosphatase, we analysed the TFIIF-containing beads for transcription activity.

Purification of the Mediator-like complex

The Mediator-like complex was isolated using conventional chromatography¹⁰ (see Supplementary Information). The complex was purified further by affinity purification using antibodies against human Med6. The complex was devoid of cdk8, but included Sur2, Med6, Med10 and cyclin C among other components of coactivator complexes.

C-terminal domain kinase assay

We carried out kinase reactions under transcription reaction conditions⁶ in the presence of all general transcription factors, DNA and γ -³²P-ATP; no other nucleotides were added to the assay. We separated products of the reaction by 5% SDS–polyacrylamide gel electrophoresis (PAGE).

Expression of cdk8/cyclin C and phosphorylation of cyclin H

Recombinant cdk8/cyclin C were induced from pET28b in the BL21 strain of *Escherichia coli* for 2 h, purified in 6 M urea, and renatured at an equimolar ratio by dialysis.

We incubated cdk8/cyclin C with purified TFIIF in the presence of γ -³²P-ATP as described²². Products of the reaction were separated by denaturing SDS–PAGE and transferred to a membrane for western blot analysis, using antibodies against cyclin H, and for autoradiography. For immunoprecipitation, we boiled the products of the kinase reaction for 5 min in 1% SDS, diluted them 10-fold, and analysed them with antibodies against cyclin H or the p44 subunit of TFIIF. We separated the products of the reactions by SDS–PAGE.

Cyclin H affinity chromatography

We carried out the assay essentially as described²⁶. Briefly, recombinant cdk8/cyclin C was passed through the affinity columns containing recombinant cyclin H or BSA (control). After extensive washes with 100 mM NaCl, Tris–HCl, pH 7.9, the columns were eluted with 100 mM glycine pH 3.5 and analysed by western blot, using the indicated antibodies.

Inducible expression of cyclin H

We induced expression of the various His-tagged cyclin H mutants was induced with 5 mM Cd²⁺ as described²⁰, and the proteins were immunopurified in the context of TFIIF with anti-ERCC3 and anti-His-tag monoclonal antibodies and subjected to further analysis. The total pool of TFIIF was first isolated from the phosphocellulose 0.5 M fraction using affinity chromatography with monoclonal antibodies against ERCC3. TFIIF was eluted from the column using a peptide²⁷ and TFIIF containing the transfected cyclin H subunit was selected in a second immunopurification step using antibodies against the His-tag. The anti-ERCC3 and anti-His-tag monoclonal antibodies were used as described^{22,27}.

Cdk-activating kinase assay

Cdk-activating kinase assay with the described mutants of TFIIF was done as described²⁸. Free CAK complexes were isolated from the phosphocellulose 0.3 M fraction using antibodies against the His-tag.

In vivo labelling

We labelled His-tagged cyclin H and its double aspartate mutant *in vivo* on induction without or with transiently coexpressed cdk8/cyclin C.

The TFIIF complex that contained the tagged cyclin H was immunopurified as above and analysed on a denaturing SDS polyacrylamide gel. The experiment was carried out in duplicates and normalized for protein and cyclin H content (data not shown). *In vivo* labelling and immunopurification were done as described²².

Cellular growth assay

We grew populations of 293EBNA cells carrying the indicated cyclin H expression vectors grown $5\text{--}10 \times 10^3$ cells per 10-cm plate in the presence of 200 nM Cd^{2+} . After two weeks, the colonies were stained with a 2% Coomassie blue, 50% methanol mix and photographed.

Received 26 April; accepted 23 June 2000.

1. Tassan, J.-P. *et al.* Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc. Natl Acad. Sci. USA* **97**, 8871–8875 (1995).
2. Li, H. *et al.* Molecular cloning and chromosomal localization of the human cyclin C (CCNC) and cyclin E (CCNE) genes: deletion of the CCNC gene in human tumors. *Genomics* **32**, 253–259 (1996).
3. Gold, M. O. *et al.* Viral transactivators E1A and VP16 interact with a large complex that is associated with CTD kinase activity and contains CDK8. *Nucleic Acids Res.* **24**, 3771–3777 (1996).
4. Rickert, P. *et al.* Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene* **12**, 2631–2640 (1996).
5. Liao, S.-M. *et al.* A kinase–cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374**, 193–196 (1995).
6. Maldonado, E. *et al.* A human RNA polymerase II complex with SRB and DNA-repair proteins. *Nature* **381**, 86–89 (1996).
7. Hengartner, C. J. *et al.* Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell* **2**, 45–53 (1998).
8. Carlson, M. Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD. *Annu. Rev. Cell. Dev. Biol.* **13**, 1–23 (1997).
9. Sun, X. *et al.* NAT a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol. Cell* **2**, 213–222 (1998).
10. Gu, W. *et al.* A novel human Srb/Med-containing cofactor complex, SMCC, involved in transcription regulation. *Mol. Cell* **3**, 97–108 (1999).
11. Boyer, T. G. *et al.* Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature* **399**, 276–279 (1999).
12. Hampsey, M. & Reinberg, D. RNA polymerase II as a control for multiple coactivator complexes. *Curr. Opin. Genet. Dev.* **9**, 132–139 (1999).
13. Andersen, G. *et al.* The structure of cyclin H: common mode of kinase activation and specific features. *EMBO J.* **16**, 958–967 (1997).
14. Rachez, C. *et al.* Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* **398**, 824–828 (1999).
15. Ryu, S., Zhou, S., Ladurner, A. G. & Tjian, R. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* **397**, 446–450 (1999).
16. Naar, A. M. *et al.* Composite coactivator ARC mediates chromatin-directed transcriptional activation. *Nature* **398**, 828–832 (1999).
17. Ge, H., Zhao, Y., Chait, B. T. & Roeder, R. G. Phosphorylation negatively regulates the function of PC4. *Proc. Natl Acad. Sci. USA* **91**, 12691–12695 (1994).
18. Kim, Y.-J. *et al.* A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599–608 (1994).
19. Gould, K. L. in *Protein Kinases* 149–176 (IRL Press, New York, 1994).
20. Kazmi, M. A. *et al.* High-level inducible expression of visual pigments in transfected cells. *Biotechniques* **21**, 304–311 (1996).
21. Hauer, C. A. *et al.* Epstein-virus episome-based promoter function in human myeloid cells. *Nucleic Acids Res.* **17**, 989–2003 (1989).
22. Akoulitch, S. & Reinberg, D. The molecular mechanism of mitotic inhibition of TFIIF is mediated by phosphorylation of CDK7. *Genes Dev.* **12**, 3541–3550 (1998).
23. Long, J. J. *et al.* Repression of TFIIF transcriptional activity and TFIIF-associated cdk7 kinase activity at mitosis. *Mol. Cell. Biol.* **18**, 1467–1476 (1998).
24. Harper, J. W. & Elledge, S. J. The role of Cdk7 in CAK function, a retro-retrospective. *Genes Dev.* **12**, 285–289 (1998).
25. Orphanides, G., Lagrange, T. & Reinberg, D. The general transcription factors of RNA polymerase II. *Genes Dev.* **10**, 2657–2683 (1996).
26. Kong, M. *et al.* Cyclin F regulates the nuclear localization of cyclin B1 through a cyclin–cyclin interaction. *EMBO J.* **19**, 1378–1388 (2000).
27. LeRoy, G., Drapkin, R., Weis, L. & Reinberg, D. Immunoaffinity purification of the human multi-subunit transcription factor IIF. *J. Biol. Chem.* **273**, 7134–7140 (1998).
28. Desai, D., Wessling, H., Fisher, R. & Morgan, D. Effects of phosphorylation by CAK on cyclin binding by CDC2 and CDK2. *Mol. Cell. Biol.* **15**, 345–350 (1995).
29. Parvin, J. D. & Sharp, P. A. DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* **73**, 533–540 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank M. Tyckocinsky for pMEP4β; R. Roeder for the cdk8–Flag cell line; E. Lees for cDNAs encoding mammalian cdk8 and cyclin C; R. Young and P. Cook for communicating unpublished results; T. Kim for affinity purified TFIIF; A. Berk for anti-hSur2 antibodies; and L. Freedman for anti-DRIP150 antibodies. This work was supported by grants from NIH and the Howard Hughes Medical Institute to D.R.

Correspondence and requests for materials should be addressed to D.R. (e-mail: reinbedf@umdj.edu).

Domain-specific recruitment of amide amino acids for protein synthesis

Debra L. Tumbula*†, Hubert D. Becker*†, Wei-zhong Chang* & Dieter Söll*‡

* Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, USA

‡ Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06511, USA

† These authors contributed equally to this work

The formation of aminoacyl-transfer RNA is a crucial step in ensuring the accuracy of protein synthesis. Despite the central importance of this process in all living organisms, it remains unknown how archaea and some bacteria synthesize Asn-tRNA and Gln-tRNA. These amide aminoacyl-tRNAs can be formed by the direct acylation of tRNA, catalysed by asparaginyl-tRNA synthetase and glutaminyl-tRNA synthetase, respectively. A separate, indirect pathway involves the formation of mis-acylated Asp-tRNA^{Asn} or Glu-tRNA^{Gln}, and the subsequent amidation of these amino acids while they are bound to tRNA, which is catalysed by amidotransferases^{1,2}. Here we show that all archaea possess an archaea-specific heterodimeric amidotransferase (encoded by *gatD* and *gatE*) for Gln-tRNA formation. However, Asn-tRNA synthesis in archaea is divergent: some archaea use asparaginyl-tRNA synthetase, whereas others use a heterotrimeric amidotransferase (encoded by the *gatA*, *gatB* and *gatC* genes). Because bacteria primarily use transamidation³, and the eukaryal cytoplasm uses glutaminyl-tRNA synthetase, it appears that the three domains use different mechanisms for Gln-tRNA synthesis; as such, this is the only known step in protein synthesis where all three domains have diverged. Closer inspection of the two amidotransferases reveals that each of them recruited a metabolic enzyme to aid its function; this provides direct evidence for a relationship between amino-acid metabolism and protein biosynthesis.

The presence in every cell of 20 aminoacyl-tRNA synthetases (AARSs), each attaching a specific amino acid directly to the cognate tRNA, has been taken for granted⁴. Previous reports^{1,5} indicated the existence of an indirect route to Gln-tRNA in *Bacillus subtilis*. The two-step pathway rests on the mis-acylation of tRNA^{Gln} with glutamate carried out by a glutamyl-tRNA synthetase

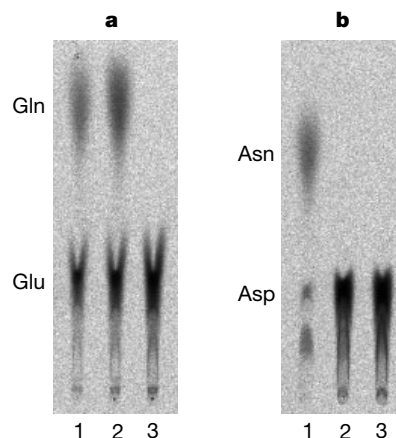


Figure 1 Transfer RNA specificity of *M. thermoautotrophicum* amidotransferases. Phosphoimages of the thin-layer chromatography separation of the tRNA-dependent conversion of ¹⁴C-labelled Glu to Gln (a) or Asp to Asn (b). Total *M. thermoautotrophicum* tRNA was used. Lane 1, Asp/Glu-AdT (GatCAB); lane 2, Glu-AdT (GatDE); lane 3, no enzyme.

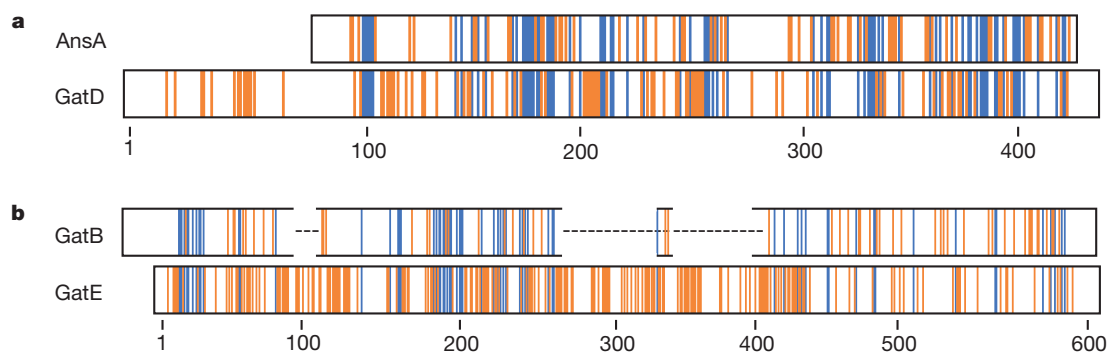
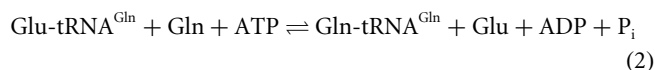
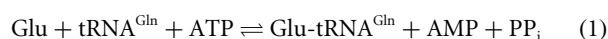


Figure 2 Conserved positions in the amino-acid alignments of amidotransferase components. **a**, AnsaA and GatD; **b**, GatB and GatE. The protein sequences listed below were aligned with CLUSTAL_X (V1.84). Blue lines represent positions with at least 75% identity between AnsaA and GatD, or between GatB and GatE. Orange lines represent the remaining positions that have at least 75% identity only within AnsaA, GatD, GatB or GatE. Numbers represent the amino-acid positions for the *M. thermoautotrophicum* GatD and GatE sequences. Gaps larger than 12 amino acids are indicated by dashed lines. AnsaA

and GatB sequences are from both the bacteria and the archaea, whereas GatD and GatE are so far from archaea only. Sequences were aligned from the following organisms: *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *M. thermoautotrophicum*, *Methanococcus jannaschii* (GatB, GatE and GatD); *Halobacterium salinarum*, *Pyrobaculum aerophilum*, *Sulfolobus solfataricus*, *Thermoplasma acidophilum* (GatE and GatD); *Pyrococcus abyssi*, *Pyrococcus furiosus*, *P. horikoshii* (GatE, AnsaA, GatD); *Aquifex aeolicus*, *Synechocystis* sp. PCC 6803, *Thermotoga maritima* (GatB); *B. subtilis* (GatB, AnsaA); and *E. coli* (AnsaA).

(GluRS) with relaxed tRNA specificity, as in equation (1).

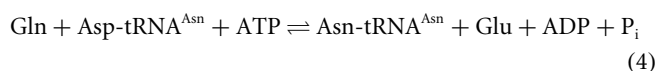
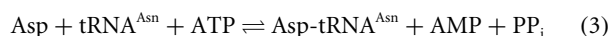


The mis-charged Glu-tRNA^{Gln} is then converted to Gln-tRNA^{Gln} by an ATP-dependent amidation of the glutamate attached to tRNA, which is catalysed by Glu-tRNA^{Gln} amidotransferase, as in equation (2). This unusual, heterotrimeric enzyme⁶ is encoded by the *gatA*, *gatB* and *gatC* genes. Biochemical and genomic data also reveal the absence of asparaginyl-tRNA synthetase (AsnRS) and glutaminyl-tRNA synthetase (GlnRS) in many organisms. Furthermore, the *Pyrococcus horikoshii* genome sequence does not suggest how Asn-tRNA and Gln-tRNA are formed. Some organisms lack recognizable aminoacyl-tRNA synthetase genes, and carry out the same function by a different enzyme (for example, class I lysyl-tRNA synthetase⁷), but this is not the case in archaea.

We searched in the completed archaeal genome sequences for candidate genes encoding an amidotransferase in open reading frames (ORFs) homologous to the *B. subtilis* *gat* genes⁶, focusing particularly on the organisms that possess a gene encoding a canonical AsnRS⁸. We identified two genes (designated *gatD* and *gatE*) that were found in all known archaeal genomes. To show that *gatDE* encodes the 'missing' amidotransferase, we cloned these genes from *Methanobacterium thermoautotrophicum* for expression in *Escherichia coli*, an organism that uses exclusively direct amino-

acylation to form Gln-tRNA. The partially purified GatDE enzyme converted Glu-tRNA^{Gln} to Gln-tRNA^{Gln} (Fig. 1a), but not Asp-tRNA^{Asn} to Asn-tRNA^{Asn} (Fig. 1b). As *gatD* and *gatE* are present in all known archaeal genomes and not in any known organism in the other two domains, we named GatDE the archaeal glutamyl-tRNA^{Gln} amidotransferase (Glu-AdT). The GatD protein (Fig. 2a) is a paralogue of *ansaA*, a gene encoding a type I L-asparaginase of unknown function⁹ found in some bacteria and archaea, whereas the GatE protein has the overall features of GatB (Fig. 2b)—one of the subunits of the bacterial amidotransferase. Both GatD and GatE are longer than their paralogue counterparts (Fig. 2); the additional sequences are strictly archaeal and show no homology to any known protein sequences. Furthermore, *gatD* and *gatE* are immediately adjacent in most archaeal genomes and possibly co-transcribed (Table 1).

Previously we showed¹⁰ that the heterotrimeric glutamyl-tRNA^{Gln} amidotransferase (encoded by *gatCAB*) from the bacteria *Deinococcus radiodurans* and *B. subtilis* can also amidate Asp-tRNA^{Asn} to Asn-tRNA^{Asn} as in equation (4), when provided with the mis-aminoacylated tRNA product of equation (3).



Searching the known archaeal genomes (Table 1) for the presence of *asnS* or *gat* genes revealed that some organisms may use direct aminoacylation but others may use the transamidation route. For

Table 1 Presence of amidotransferase and AsnRS genes in archaea

Organism	GatD*	GatE*	Adjacency (<i>gatD</i> and <i>gatE</i>)	AsnRS†	Asp/Glu-AdT‡
<i>Aeropyrum pernix</i>	427	642	Yes	—	+
<i>Archaeoglobus fulgidus</i>	408	613	No†	—	+
<i>Halobacterium salinarum</i>	427	622	No	—	+
<i>Methanobacterium thermoautotrophicum</i>	435	619	Yes	—	+
<i>Methanococcus jannaschii</i>	417	630	Yes	—	+
<i>Methanosarcina mazei</i>	421	632	No	—	+
<i>Sulfolobus solfataricus</i>	444	633	Yes	—	+
<i>Sulfolobus tokodaii</i>	448	628	Yes	—	+
<i>Pyrobaculum aerophilum</i>	417	608	No	+	—
<i>Pyrococcus abyssi</i>	438	633	Yes	+	—
<i>Pyrococcus furiosus</i>	438	628	Yes	+\$	—
<i>Pyrococcus horikoshii</i>	438	632	Yes	+	—
<i>Thermoplasma acidophilum</i>	409	603	Yes	+	—

* The numbers represent the predicted number of amino acids in each subunit. In contrast to GatE, the average length of the GatB sequences in the alignment of Fig. 2 is 482 amino acids.

† Presence (+) or absence (—) of the encoding ORF.

‡ A truncated *gatD* gene (AF1441) is adjacent to *gatE* (AF1440), but an intact *gatD* gene (AF0882) is present elsewhere.

\$ The enzyme was shown to be active in a cell extract.

example, *M. thermoautotrophicum* possesses orthologues of the *gatC*, *gatA* and *gatB* genes. We therefore cloned these genes next to each other into an *E. coli* vector. The overexpressed GatCAB enzyme possessed dual tRNA specificity: in addition to Asn-tRNA formation, it also generated Gln-tRNA (Fig. 1). Thus, we designate this enzyme Asp/Glu-AdT. The presence of these two different amidotransferase enzymes (Glu-AdT and Asp/Glu-AdT) accounts for the two separable activities observed in *M. thermoautotrophicum* cell extracts (data not shown). The existence of two amidotransferases and the interpretation of the genomic data (Table 1) show that the archaea divide into two classes regarding the synthesis of amide aminoacyl-tRNAs. As the *gatDE*-encoded Glu-AdT is found in all 13 archaea with known genome sequences (Table 1), it is probably responsible for Gln-tRNA formation. Asn-tRNA synthesis in archaea is not, however, uniform (Table 1). Eight of the known archaea (*A. pernix*, *A. fulgidus*, *H. salinarum*, *M. thermoautotrophicum*, *M. jannaschii*, *M. mazei* and two *Sulfolobus* species) do not contain *asnS* genes and must use the *gatCAB*-encoded Asp/Glu-AdT for Asn-tRNA formation; but *P. aerophilum*, *T. acidophilum* and *Pyrococcus* spp. do use AsnRS, as they carry *asnS* and lack *gatCAB* genes. In these archaea, therefore, GatDE is presumably the sole source of Glu-AdT activity.

The existence of seemingly redundant Glu-AdT activities in many archaea is intriguing. Given the presence of the *gatCAB* genes encoding Asp/Glu-AdT in some archaeal genomes, why does every archaeal genome still carry a *gatDE*-encoded Glu-AdT? Both GatCAB and GatDE recognize the same tRNA^{Gln} substrate (Fig. 1); the *in vivo* nitrogen amide donor is not known. An examination of potential substrates indicated that Glu-AdT has no substrate advantage over the archaeal GatCAB enzyme. Asp/Glu-AdT is able to use glutamine, asparagine or ammonium chloride as the donors (Fig. 3a), whereas Glu-AdT cannot use ammonium (Fig. 3b). Possibly, the enzymes are involved in different metabolic functions.

AsnRS is found in almost all bacteria (exceptions are *Chlamydia trachomatis*, *Helicobacter pylori*, *Aquifex aeolicus*, *Mycobacterium tuberculosis* and *Mycobacterium leprae*), presumably as a transfer from eukarya^{8,11}. GlnRS has so far only been found in the *Thermus/Deinococcus* group and some members of the proteobacteria^{3,10,12–14}. Sequence analyses support the case of horizontal transfer of eukaryal GlnRS after these domains were established^{14–16}. The bacterial Asp/Glu-AdT encoded by the *gatCAB* genes is capable of synthesizing both Asn-tRNA and Gln-tRNA¹⁰. The enzyme reaction in any cell is dependent on the biochemical context provided by the availability of the mis-acylated substrate to be amidated. For example, *B. subtilis*⁶ contains a non-discriminating GluRS capable of making Glu-tRNA^{Gln}, but not a mis-charging aspartyl-tRNA synthetase (AspRS) for mis-acylating tRNA^{Asn}; thus, in *B. subtilis* the Asp/Glu-AdT carries out Gln-tRNA formation.

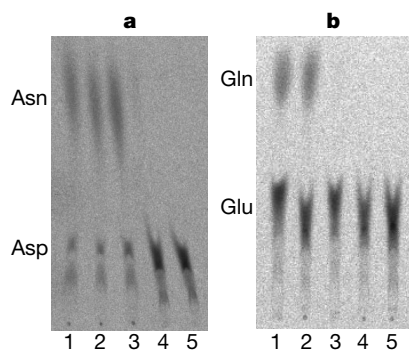


Figure 3 Effect of different amide donors (2 mM) on amidation activity. Lane 1, glutamine; lane 2, asparagine; lane 3, ammonium chloride; lane 4, no addition; lane 5, all donors in the absence of enzyme. **a**, Asp/Glu-AdT (GatCAB). **b**, Glu-AdT (GatDE).

In contrast, *D. radiodurans* and *Thermus thermophilus* produce Asp-tRNA^{Asn} by a second, archaeal-type AspRS, but possess a discriminating GluRS^{10,12}; thus, in these organisms the Asp/Glu-AdT is specific for Asn-tRNA formation. Preliminary data show that Asp/Glu-AdT enzymes in *H. pylori* (D.L.T., K. Melchers and A. Buhmann, unpublished data) and *C. trachomatis* (G. Racznik, A. Miller and H.D.B., unpublished data) make both Asn-tRNA and Gln-tRNA. Thus, a key factor maintaining the indirect transamidation route is a mis-acylating aminoacyl-tRNA synthetase. The Asp/Glu-AdT enzyme must still be capable of highly specific RNA recognition, however, as it will not use the cognate Asp-tRNA^{Asp} or Glu-tRNA^{Glu} as substrates.

All available data show that the eukaryal cytoplasm uses AsnRS and GlnRS to make Asn-tRNA and Gln-tRNA, respectively. Animal and yeast mitochondria and plant chloroplasts are expected to carry a *gatA/gatB*-based amidotransferase to make Gln-tRNA^{13,17}, in keeping with the concept of the bacterial origin of these organelles. The only demonstrated exception in the organelles is a mitochondrial GlnRS found in the deeply rooted trypanosomatids¹⁸.

Phylogenetic and sequence analysis of genome data supported by biochemical studies have pointed to widespread horizontal transfer of aminoacyl-tRNA synthetase genes between the different domains in the living world^{8,11}. Thus, it was surprising to see that *glnS* was not transferred from its eukaryal origin to the archaea^{2,5,19}. We therefore considered whether the preponderance of indirect pathways to Gln-tRNA in archaea may be mostly due to differences in the crucial identity elements of the tRNAs. For example, *E. coli* GlnRS cannot aminoacylate unfractionated tRNA from *M. thermoautotrophicum* (Fig. 4a). Although yeast GlnRS cannot aminoacylate bacterial tRNA²², the yeast enzyme also could not glutaminylate total *M. thermoautotrophicum* tRNA (data not shown). The principal sequence elements of *E. coli* tRNA^{Gln} responsible for its recognition by *E. coli* GlnRS are well known at both the biochemical and the structural level^{20,21}. *M. thermoautotrophicum* has two tRNA^{Gln} species (tRNA^{Gln1} and tRNA^{Gln2}) with different bases at over half of the critical identity element positions when compared to bacterial

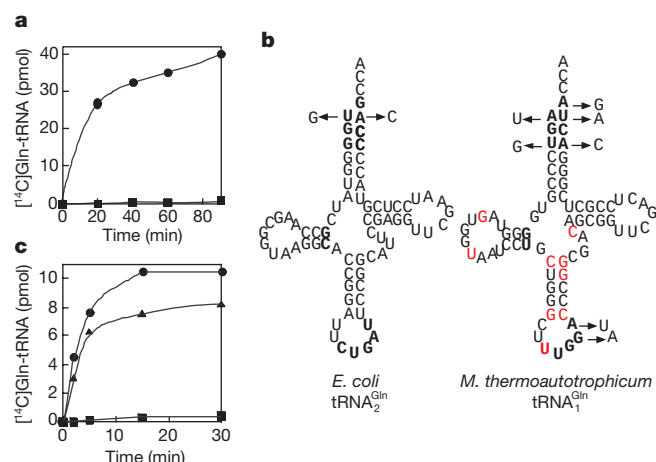


Figure 4 Transfer RNA specificity of *E. coli* GlnRS. Aminoacylation was carried out as described in Methods. **a**, Unfractionated *E. coli* tRNA (circles) and total *M. thermoautotrophicum* tRNA (squares). **b**, Cloverleaf structures of the tRNAs. Arrows indicate nucleotide changes in the *E. coli* tRNA^{Gln1} transcript and in the *M. thermoautotrophicum* mutant tRNA^{Gln2} transcript. The change of the first base pair in the *E. coli* tRNA^{Gln2} transcript has negligible effect on aminoacylation²⁰. Shadowed nucleotides are major identity elements for *E. coli* GlnRS²¹. Nucleotides in red differ between *M. thermoautotrophicum* tRNA^{Gln1} and tRNA^{Gln2}. **c**, *E. coli* tRNA^{Gln2} transcript (circles), *M. thermoautotrophicum* tRNA^{Gln1} transcript (squares), and *M. thermoautotrophicum* mutant tRNA^{Gln1} transcript (triangles). Wild-type *M. thermoautotrophicum* tRNA^{Gln1} transcript could be aminoacylated with glutamate using GluRS from *M. thermoautotrophicum* (data not shown).

tRNA^{Gln} (Fig. 4b). Similar to the results with total tRNA, *E. coli* GlnRS did not charge a transcript of tRNA^{Gln} from *M. thermoautotrophicum* (Fig. 4c); however, 'transplanting' the principal *E. coli* tRNA^{Gln} identity elements into this transcript (Fig. 4b) allowed efficient glutaminylation close to that of the *E. coli* tRNA^{Gln} transcript (Fig. 4c). Thus, the archaeal tRNA^{Gln} sequence itself may have prevented the stable transfer of GlnRS from the other domains into archaea. Alternatively, this archaeal sequence may have evolved later as a consequence of different RNA-protein interactions required by the different Gln-tRNA synthetic route.

Our work completes the picture of the biosynthetic pathways for amide aminoacyl-tRNA formation in the living world. Most unexpected is that the three domains use different routes and enzymes. This is the only known step in protein synthesis where all three domains diverge. Transamidation is a very plausible and utilitarian route, which incorporates the mechanistic concept of Asn and Gln biosynthesis into that of aminoacyl-tRNA formation. The evolution of the archaeal Glu-AdT encompassed the recruitment (into *gatD*) of the amino-acid-metabolizing enzyme asparaginase, whereas the *gatA* subunit appears to be related to amidases involved in general amide bond cleavage²³. Furthermore, the tRNA-dependent transamidation pathway may be the sole biosynthetic route to asparagine in the *Thermus/Deinococcus* group^{10,12}. These examples implicate Asp and Glu as the earliest amino acids and support the idea that amino-acid metabolism had an important role in the evolution of the protein synthesis system²⁴. Alternatively, the synthesis of some amino acids might have evolved after the establishment of protein synthesis to provide another supply route of amino acids or a new amino acid for translation. This may explain the involvement of a histidyl-tRNA synthetase paralogue in histidine biosynthesis²⁵. Thus, there might have been a closer relationship between aminoacyl-tRNA synthetases and amino-acid biosynthetic enzymes than previously anticipated²⁶.

The domain distributions of the amide aminoacyl-tRNA pathways raise other evolutionary questions. Are the transamidation routes relics of mechanisms that led to expansion of the genetic code²⁷? Why have horizontal transfers of the two synthetases not led to the elimination of the amidotransferases in modern organisms? Possibly these multi-component complexes (of an RNA binding protein with a metabolic enzyme) evolved earlier, in part driven by the metabolic conditions of the organism. Despite widespread lateral transfer of AsnRS and GlnRS^{8,11,14–16}, however, transamidation may not be easily replaced by a single gene transfer because of its many enzymes and its possible requirement for amino-acid metabolism^{10,12}. The fact that tRNA identity may have been important in preventing a stable GlnRS transfer into the archaea is appealing. Similarly, a lateral gene transfer may have been 'fixed' to correct a poor protein-tRNA interaction for the class I lysyl-tRNA synthetase transfer into spirochetes²⁸. This new understanding of Asn-tRNA and Gln-tRNA formation and the demonstration of dual-specificity aminoacyl-tRNA synthetases²⁹ expands the accepted view of aminoacyl-tRNA synthesis and opens a new chapter in the investigation of the translation system and its origins. □

Methods

Plasmid constructs

Genes were amplified by polymerase chain reaction from *M. thermoautotrophicum* strain ΔH genomic DNA, cloned into pET20b (Novagen) and expressed in *E. coli* BL21(DE3) containing the minor-tRNA overproducing vector pSJS1240 (ref. 30). The *gatA*, *gatB* and *gatC* genes were cloned together in a *B. subtilis* operon-like arrangement⁶ as was done for *D. radiodurans*¹⁰. The intergenic sequence inserted between *gatC* and *gatA* and also between *gatA* and *gatB* was 5'-AGAGGATATATATAT-3' (ref. 10). We purified *M. thermoautotrophicum* GluRS by fusing the gene at its 3' end to the intein gene in the vector pCYB1 (New England Biolabs), followed by heterologous expression in *E. coli* BL21(DE3)/pSJS1240, affinity chromatography and cleavage.

Preparation of heterologously expressed amidotransferases

After expressing the *M. thermoautotrophicum* genes in *E. coli* BL21(DE3)/pSJS1240, we collected and lysed the cells (32 g for *gatCAB*, and 14 g for *gatDE*) by sonication in buffer A

(50 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 1 mM benzamidinium-HCl and 0.1 mM Na₂EDTA). After a 15-min incubation at 65 °C, thermolabile proteins were pelleted by centrifugation at 8,700g for 20 min at 4 °C. The solution was filtered (Gelman Acrodisc, 0.45 μm) and applied at 1 ml min⁻¹ to a Q Sepharose column (Pharmacia, 11 cm × 3 cm²) equilibrated in buffer A. After washing with a further 120 ml of buffer A, elution was by a 375 ml gradient (0–0.75 M KCl) at 3 ml min⁻¹. Active fractions eluted at 0.5–0.6 M KCl. At this point, *GatDE* (5 mg total protein) was over 50% pure, as judged by SDS-polyacrylamide gel electrophoresis. *GatCAB* was similarly purified to about 2% purity. Active fractions eluted at 0.45–0.55 M KCl. Both enzyme preparations were dialysed against buffer A, glycerol was added to 50% (v/v) and enzymes were stored at –20 °C.

Preparation of tRNAs

We extracted total tRNA from frozen cells of *M. thermoautotrophicum* strain Marburg as described¹⁰ including a sonication step before phenol extraction. Transcripts of *M. thermoautotrophicum* tRNA^{Gln} genes were made from 5' cis-acting hammerhead ribozyme constructs using *Bst*NI digestion of the pUC18 clones to provide the 3' terminus. We carried out transcription reactions at 37 °C for 4 h in 40 mM Tris-HCl (pH 8.0), 22 mM MgCl₂, 2 mM spermidine-HCl, 10 mM DTT, nucleoside triphosphates (4 mM each), 50 μg ml⁻¹ BSA, 10 μg ml⁻¹ pyrophosphatase, 0.06 μg ml⁻¹ digested vector DNA and 50 μg ml⁻¹ T7 RNA polymerase. Cleavage was facilitated by fivefold dilution of the RNA solution into 30 mM MgCl₂ and incubation for 1 h at 60 °C. After phenol-chloroform extraction and ethanol precipitation, we purified the RNA by denaturing polyacrylamide gel electrophoresis, two more phenol-chloroform extractions and ethanol precipitation. The RNA transcripts were denatured at 80 °C for 5 min in 2 mM MgCl₂ and allowed to renature by gradual cooling.

Amidotransferase assays

We prepared mis-aminoacylated tRNAs at 37 °C with purified GluRS from *M. thermoautotrophicum* or AspRS2 from *D. radiodurans*¹⁰ and isolated them as described¹⁰. Reactions contained 4 μM aminoacyl-tRNA synthetase, 5–12 μM tRNA transcript or 110–130 μM total tRNA and 60 μM [¹⁴C]aspartate (213 mCi mmol⁻¹, 50 μCi ml⁻¹) or [¹⁴C]glutamate (260 mCi mmol⁻¹, 50 μCi ml⁻¹). We carried out amidotransferase reactions at 37 °C as described¹⁰, but included glutamine, asparagine and ammonium chloride (2 mM of each), 4,000–10,000 c.p.m. of [¹⁴C]aminoacyl-tRNA and 18 or 0.5 μg of the *GatCAB* or *GatDE* preparations, respectively. To assay^{6,10} the conversion, aminoacyl-tRNA was extracted, precipitated and deacylated, and the resulting free amino acids were separated by TLC in chloroform:methanol:ammonium hydroxide:water (6:6:2:1)⁶. The [¹⁴C]amino acids were detected by phosphorimaging and their positions verified by ninhydrin visualization of unlabelled standards.

Glutaminyl-tRNA synthetase assay

GlnRS assay reactions (50 μl) contained 50 mM HEPES-NaOH, pH 7.2, 25 mM KCl, 10 mM MgCl₂, 4 mM ATP, 5 mM dithiothreitol, 125 μM [¹⁴C]glutamine (257 mCi mmol⁻¹, 50 μCi ml⁻¹), 2.3 μM *E. coli* GlnRS and 1 μM transcript or 33 μM total tRNA. We spotted 10-μl aliquots onto Whatman 3MM filter discs and measured TCA-precipitable radioactivity as described¹⁰.

Received 3 April; accepted 22 June 2000.

1. Wilcox, M. & Nirenberg, M. Transfer RNA as a cofactor coupling synthesis with that of protein. *Proc. Natl Acad. Sci. USA* **61**, 229–236 (1968).
2. Curnow, A. W., Ibba, M. & Söll, D. tRNA-dependent asparagine formation. *Nature* **382**, 589–590 (1996).
3. Gagnon, Y., Lacoste, L., Champagne, N. & Lapointe, J. Widespread use of the Glu-tRNA^{Gln} transamidation pathway among bacteria. *J. Biol. Chem.* **271**, 14856–14863 (1996).
4. Söll, D. & Schimmel, P. R. in *The Enzymes* Vol. 10 (ed. Boyer, P.) 489–538 (Academic, New York, 1974).
5. Bayley, S. T. & Griffiths, E. Codon assignments and fidelity of translation in a cell-free protein-synthesizing system from an extremely halophilic bacterium. *Can. J. Biochem.* **46**, 937–944 (1967).
6. Curnow, A. W. *et al.* Glu-tRNA^{Gln} amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc. Natl Acad. Sci. USA* **94**, 11819–11826 (1997).
7. Ibba, M. *et al.* A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases. *Science* **278**, 1119–1122 (1997).
8. Woese, C. R., Olsen, G., Ibba, M. & Söll, D. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* **64**, 202–236 (2000).
9. Spring, K. J., Jerlström, P. G., Burns, D. M. & Beacham, I. R. L-asparaginase genes in *Escherichia coli*: isolation of mutants and characterization of the *ansA* gene and its protein product. *J. Bacteriol.* **166**, 135–142 (1986).
10. Curnow, A. W., Tumbula, D. L., Pelaschier, J. T., Min, B. & Söll, D. Glutamyl-tRNA^{Gln} amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proc. Natl Acad. Sci. USA* **95**, 12838–12843 (1998).
11. Wolf, Y. I., Avirand, L., Grishin, N. V. & Koonin, E. V. Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**, 689–710 (1999).
12. Becker, H. D. & Kern, D. *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc. Natl Acad. Sci. USA* **95**, 12832–12837 (1998).
13. Schön, A., Kannangara, C. G., Gough, S. & Söll, D. Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* **331**, 187–190 (1988).
14. Brown, J. R. & Doolittle, W. F. Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. *J. Mol. Evol.* **49**, 485–495 (1999).
15. Handy, J. & Doolittle, R. F. An attempt to pinpoint the phylogenetic introduction of glutaminyl-tRNA synthetase among bacteria. *J. Mol. Evol.* **49**, 709–715 (1999).

16. Siatecka, M., Rozek, M., Barciszewski, J. & Mirande, M. Modular evolution of the Glx-tRNA synthetase family: rooting of the evolutionary tree between the bacteria and archaea/eukarya branches. *Eur. J. Biochem.* **256**, 80–87 (1998).
17. Kim, S.-I. *et al.* A nuclear genetic lesion affecting *Saccharomyces cerevisiae* mitochondrial translation is complemented by a homologous *Bacillus* gene. *J. Bacteriol.* **179**, 5625–5627 (1997).
18. Nabholz, C. E., Hauser, R. & Schneider, A. *Leishmania tarentolae* contains distinct cytosolic and mitochondrial glutamyl-tRNA synthetase activities. *Proc. Natl Acad. Sci. USA* **94**, 7903–7908 (1997).
19. Gupta, R. *Halobacterium volcanii* tRNAs: identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. *J. Biol. Chem.* **259**, 9461–9471 (1984).
20. Ibba, M., Hong, K. W., Sherman, J. M., Sever, S. & Söll, D. Interactions between tRNA identity nucleotides and their recognition sites in glutamyl-tRNA synthetase determine the cognate amino acid affinity of the enzyme. *Proc. Natl Acad. Sci. USA* **93**, 6953–6958 (1996).
21. Hayase, Y. *et al.* Recognition of bases in *Escherichia coli* tRNA^{Gln} by glutamyl-tRNA synthetase: a complete identity set. *EMBO J.* **11**, 4159–4165 (1992).
22. Ludmerer, S. W. & Schimmel, P. Gene for yeast glutamine tRNA synthetase encodes a large amino-terminal extension and provides a strong confirmation of the signature sequence for a group of the aminoacyl-tRNA synthetases. *J. Biol. Chem.* **262**, 10801–10806 (1987).
23. Fourmand, D., Bigey, F. & Arnaud, A. Acyl transfer activity of an amidase from *Rhodococcus* sp. strain R312: formation of a wide range of hydroxamic acids. *Appl. Environ. Microbiol.* **64**, 2844–2852 (1998).
24. Wong, J. T.-F. A co-evolution theory of the genetic code. *Proc. Natl Acad. Sci. USA* **72**, 1909–1912 (1975).
25. Sissler, M. *et al.* An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc. Natl Acad. Sci. USA* **96**, 8985–8990 (1999).
26. Weiner, A. M. Molecular evolution: aminoacyl-tRNA synthetases on the loose. *Curr. Biol.* **9**, R842–R844 (1999).
27. Di Giulio, M. The RNA world, the genetic code and the tRNA molecule. *Trends Genet.* **16**, 17–19 (2000).
28. Ibba, M., Bono, J. L., Rosa, P. A. & Söll, D. Archaeal-type lysyl-tRNA synthetase in the Lyme disease spirochete *Borrelia burgdorferi*. *Proc. Natl Acad. Sci. USA* **94**, 14383–14388 (1997).
29. Stathopoulos, C. *et al.* One polypeptide with two aminoacyl-tRNA synthetase activities. *Science* **287**, 479–482 (2000).
30. Kim, R. *et al.* Overexpression of archaeal proteins in *Escherichia coli*. *Biotechnol. Lett.* **20**, 207–210 (1998).

Acknowledgements

We thank R. Hedderich for *M. thermoautotrophicum* Marburg cells, J. Reeve for *M. thermoautotrophicum* ΔH DNA, K. O. Stetter for *Pyrococcus* cells, and H. Kobayashi for *E. coli* GlnRS and *E. coli* tRNA^{Gln} transcript. We also thank M. Ibba for critically reading the manuscript and S. Fitz-Gibbon, T. Hartsch, A. Johann, D. Oesterhelt, A. Ruepp and S. Schuster for sharing unpublished sequence data. D.L.T. and H.D.B. are postdoctoral fellows of the National Institute of General Medical Sciences and the EMBO, respectively. This work was supported by grants from the National Institute of General Medical Sciences.

Correspondence and requests for materials should be addressed to D.S.
(e-mail: soll@trna.chem.yale.edu).

erratum

Atomically defined mechanism for proton transfer to a buried redox centre in a protein

Kaisheng Chen, Judy Hirst, Raul Camba, Christopher A. Bonagura, C. David Stout, Barbara K. Burgess & Fraser A. Armstrong

Nature **405**, 814–817 (2000).

The title to Box 1 in this paper was printed incorrectly. The correct title is 'Sequential electron and proton transfers at the [3Fe-4S] cluster in ferredoxin I'. □

corrections

The duration of antigen receptor signalling determines CD4⁺ versus CD8⁺ T-cell lineage fate

Koji Yasutomo, Carolyn Doyle, Lucio Miele, Chana Fuchs & Ronald N. Germain

Nature **404**, 506–510 (2000).

Chana Fuchs was inadvertently omitted from the list of authors. Dr Fuchs is affiliated with the Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892, USA. □

The DNA sequence of human chromosome 21

The chromosome 21 mapping and sequencing consortium

M. Hattori, A. Fujiyama, T. D. Taylor, H. Watanabe, T. Yada, H.-S. Park, A. Toyoda, K. Ishii, Y. Totoki, D.-K. Choi, Y. Groner, E. Soeda, M. Ohki, T. Takagi, Y. Sakaki; S. Taudien, K. Blechschmidt, A. Polley, U. Menzel, J. Delabar, K. Kumpf, R. Lehmann, D. Patterson, K. Reichwald, A. Rump, M. Schillhabel, A. Schudy, W. Zimmermann, A. Rosenthal; J. Kudoh, K. Schibuya, K. Kawasaki, S. Asakawa, A. Shintani, T. Sasaki, K. Nagamine, S. Mitsuyama, S. E. Antonarakis, S. Minoshima, N. Shimizu; G. Nordsiek, K. Hornischer, P. Brant, M. Scharfe, O. Schön, A. Desario, J. Reichelt, G. Kauer, H. Blöcker; J. Ramser, A. Beck, S. Klages, S. Hennig, L. Riesselmann, E. Dagand, T. Haaf, S. Wehrmeyer, K. Borzym, K. Gardiner, D. Nizetic, F. Francis, H. Lehrach, R. Reinhardt & M.-L. Yaspo

Nature **405**, 311–319 (2000).

The name of one author, Y. Groner, was inadvertently omitted from the author list but is included above. He is affiliated with the Department of Molecular Genetics and Human Genome Center, The Weizmann Institute of Science, Rehovot, Israel. Also, in the second paragraph of the section headed 'Comparison of chromosome 21 with other autosomes', the second sequence should read 'For example, a 100-kb region of a contig (AP001656) on 21p is shared with chromosomes 4, 7, 20 and 22.' □

Proteins, peptides, proteomics

A 10-fold boost for a proteomics package, and new ways with detection.

Hyperfilm ECL and MP

Amersham Pharmacia Biotech

www.apbiotech.com

Increased sensitivity autoradiography films for protein and nucleic acid blot detection

The double-coated emulsion in Hyperfilm ECL aims for improved clarity and sensitivity with both blue and green light chemiluminescent systems. The film can be processed in automatic processors or manually, using most common X-ray film developers and fixers. It is suitable for ECL, ECL Plus, CDP-Star and all other chemiluminescent systems. Hyperfilm MP is a multipurpose film designed for use in direct and indirect autoradiography. It is suited to detection of ^{32}P , ^{35}P and ^{125}I southern, northern and western blots, ^{35}S and ^{32}P sequencing and fluorography of ^3H , ^{14}C and ^{35}S . New formulation Hyperfilm MP is also available in an envelope sealed format giving maximum protection and ease of use in the life science laboratory.

Reader Service No. 100

96-well Equilibrium Dialyser

Amika Corp.

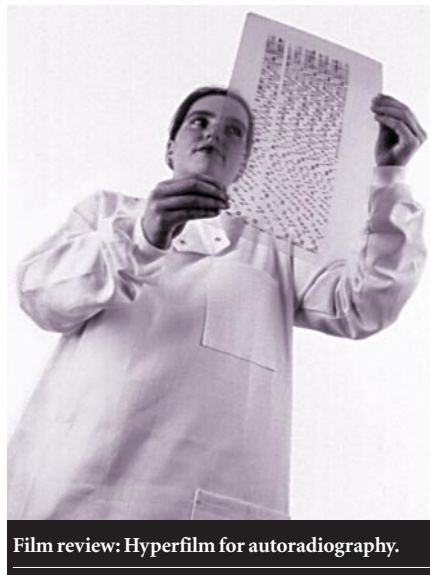
www.amika.com

Equilibrium dialysis simultaneously on 96 micro-volume samples.

This new dialyser is designed for the study of the binding of small molecules or ions to macromolecules, such as DNA, proteins, peptides, carbohydrates and lipids. The 96-well Equilibrium Dialyser offers precisely standardized conditions for up to 96 wells by ensuring uniformity of membrane area, sample



Amika's handy new dialysis system.



Film review: Hyperfilm for autoradiography.

volume, degree of agitation and temperature. The volume range for each sample is 100–200 μl per well. The system is available with membranes with MWCOs of 5,000 or 10,000 daltons. Amika also provides a Plate Rotator, which can rotate the plate 360° at a uniform speed to ensure sufficient interaction between samples across the membrane. The 96-well Equilibrium Dialyser is also well suited for high-throughput screening applications.

Reader Service No. 101

Proteins A and G

KPL

www.kpl.com

Immunoprecipitation kits for the purification of Protein A and G

KPL's new Immunoprecipitation Kits offer a rapid and convenient means for purification of a specific antigen from a complex mixture of proteins. Both Protein A and Protein G kits are available, comprising a buffer system, sodium chloride and detergents suitable for the preparation of lysis and wash solutions for a broad range of systems. The kits also include protease inhibitor cocktail to minimize the loss of target protein due to protease activity during lysis. Following electrophoresis and transfer, KPL's western blotting kits can be used for chemiluminescent or chromogenic detection.

Reader Service No. 102

BlueRanger

Pierce

www.piercenet.com

Molecular weight markers ready to go

The BlueRanger Prestained Protein Molecular Weight Marker Mix eliminates the need to

expose protein molecular weight markers to freeze-thaw cycles and its single-dose packaging eliminates the risk of contamination due to multiple withdrawals. This individually dried and stabilized mixture of seven proteins spans the range 18.3–250 kD. Each well contains sufficient dried, prestained proteins for loading one or two SDS-PAGE gel lanes. To use the 48-well microtube plate, simply puncture the foil covering a single well containing the dried marker mix with a pipette tip and add 10 μl deionized water. The marker proteins are reconstituted instantly and ready for loading onto a gel lane. BlueRanger Marker is compatible with a broad range of SDS-PAGE gels and downstream applications.

Reader Service No. 103

TurboSEQUEST

ThermoQuest

www.thermoquest.com

More information on a proteomics software package

A new brochure has been released covering the use of TurboSEQUEST software for automatic protein identification. The package compares and correlates uninterpreted tandem (MS/MS) mass spectra of peptides with theoretical MS/MS spectra generated from protein or nucleotide databases. Its cross-correlation algorithm is claimed to be especially suited to identifying proteins of relatively low abundance in protein mixtures. TurboSEQUEST is the successor to — and is about ten times faster than — the SEQUEST package.

Reader Service No. 104

Protein detection kits

Pall Gelman

www.pall.com/gelman

Protein detection without radiolabels

Pall Gelman Laboratory's new Protein Detection Kits with BioTrace PVDF Membrane are designed for the sensitive detection of proteins using horseradish peroxidase labelled secondary antibodies and LumiGLO Chemiluminescent Substrate. The kits include membrane and reagents developed to perform optimally together and are claimed to



β -Gal detection, using Pall Gelman's new kit.

new on the market

give sensitivity equal to that of radioactive detection. Proteins can be detected after electrophoresis and transfer (western blotting) or quantitated by dot blotting directly on the membrane. The new kits use an enzyme-linked affinity purified secondary antibody coupled to horseradish peroxidase and a sensitive chemiluminescent substrate. Signal can be measured using a camera luminometer or by exposing X-ray film to the membrane. If necessary, blots can be stripped and reprobed after detection.

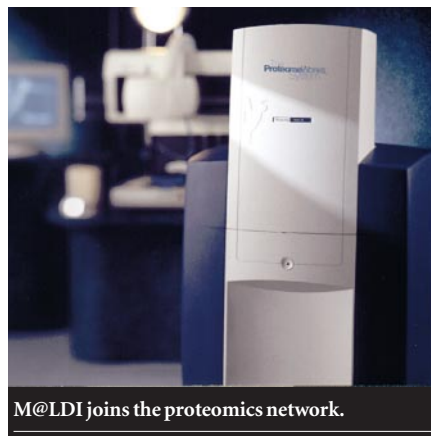
Reader Service No. 105

M@LDI

Micromass www.micromass.co.uk

The 2D gel MS analyser of the Proteome Works system

M@LDI is a networked 2D-Gel-MS analyser designed for the high-throughput protein identification techniques developed for proteomic research. The system automates the acquisition of optimized mass spectra from in-gel protein digests; monoisotopic peptide mass fingerprint (mPMF) determination; interrogation of multiple FASTA formatted databases; e-reporting of protein assign-



M@LDI joins the proteomics network.

ments; and auto-annotation of PDQuest gel images with protein ID. The M@LDI Application-Manager is scalable allowing automated assignment of protein identity on either a stand-alone PC or a corporate client-server network. Networking enables distribution of the data capture, protein assignment and results presentation functions, within a secure client-server architecture, thus maximizing computing power.

Reader Service No. 106

Glucocorticoid receptor

PanVera Corporation www.panvera.com

A receptor for steroid hormone binding studies

A new recombinant human glucocorticoid receptor (GR) has been added to the PanVera range. This receptor is a 94 kDa ligand-activated intracellular transcriptional regulator and is the mediator of glucocorticoid action that regulates the metabolism of carbohydrates, proteins and fats, suppresses the immune/inflammatory responses, activates the central nervous system, regulates cardiovascular function, and affects basal and stress-related homeostasis. PanVera's GR is a full-length, human, steroid hormone receptor, produced from recombinant baculovirus in insect cells and is likely to be associated with heat shock and other endogenous proteins. The preparation is partially purified in order to maintain high activity.

Reader Service No. 107

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound elsewhere in the journal.

ADVERTISEMENTS



**Cross species
Primer design
with
Primer Premier**

- ✓ Allele specific primers
- ✓ Degenerate Primers
- ✓ ClustalW alignment
- ✓ Pathogenic primers
- ✓ Restriction enzyme & motif analysis
- ✓ Nested/multiplex primers and more...

Download free demo from
<http://www.PremierBiosoft.com>

Call 888-847-7494 (Toll free in US)
650-856-2703 (International)

PREMIER
Biosoft International
Intuitive software for the molecular biologist

READER ENQUIRY NO. 45

Diagnostic Reagents

Pefachrome®-Series: chromogenic and fluorigenic substrates

Pefabloc®-Series: serine proteinase inhibitors

Purified snake venom proteins (e.g. Protac®, Ecarin, Batroxobin)

PENTAPHARM LTD
CH 4002 Basel Fax: +41 61 319 96 19
<http://www.pentapharm.ch>

READER ENQUIRY NO. 16

Database Management Systems

Bioinformatics
Powered by GCG

- Comprehensive analysis packages
 - Enterprise-wide solutions
 - Extendable software
 - Computing security
 - Scriptable analysis

www.gcg.com

GCG is an Oxford Molecular Company

Client/Server & Desktop Software

READER ENQUIRY NO. 36



**Look no further
for the best value in
Oligo Synthesis**

Competitive pricing
Routine 24 hour dispatch • Every Oligo QC verified by PAGE
Rigorous batch QC by PCR and Mass Spec
Web ordering and tracking • Local Freephone numbers
A great deal in any language!!

SIGMA GENOSYS (+44) (0) 1223 839000 • fax (+44) (0) 1223 839200
email: info@sigma-genosys.co.uk • www.sigma-genosys.co.uk

READER ENQUIRY NO. 33

**Lysyl Endopeptidase
(Achromobacter Protease 1)**
Specifically cleaves all Lys-X bonds including Lys-Pro.
Highly Stable, Water Soluble



Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 23

Custom Peptides & Polyclonal Antibodies

Please contact us about complete services for your custom peptide and polyclonal antibody needs.

Research Genetics
U.S. or Canada 800-533-4363
FAX 256-536-9016

Check out our homepage at
<http://www.resgen.com>

READER ENQUIRY NO. 39

Reversal of fortune

Influx of 'new' money is revitalizing Canadian science

p113**Healthy investment**

Reorganization breathes new life into health research in Canada

p113

Research gets a new lease of life as Canada takes the initiative

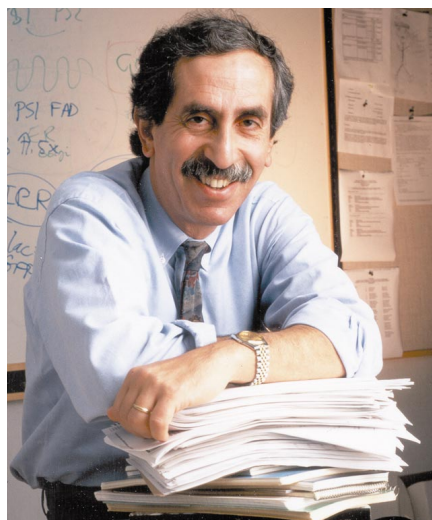
After years of cuts, Canadian science is finally receiving the financial backing it needs to flourish in the 21st century, says Diane Gershon.

By any measure, these are the best times that researchers in Canada have seen for many years. New federal and provincial initiatives are rejuvenating all aspects of the research enterprise. The federal government is "serious about positioning Canada as a place of choice to pursue academic research careers", says Robert Davidson, director of research and policy analysis at the Association of Universities and Colleges of Canada (AUCC).

The new initiatives provide a "great signal to young Canadians that there's a future in the country for people wanting to do research", says Alan Bernstein, president of the newly formed Canadian Institutes of Health Research (CIHR, see below).

This turnaround is due in part to "the realization that something had to be done if Canada was going to be competitive in this knowledge economy", says David Strangway, president and chief executive officer of the Canada Foundation for Innovation (CFI). The CFI is an independent agency set up by the government in 1997 to fund infrastructure development in Canada's research institutions.

The situation today is a far cry from that faced by researchers in Canada for much of the past decade. When the present government took office in November 1993, its initial focus was on deficit reduction, and no one was spared in its fiscal belt-tightening



Bernstein: strengthening health research.

measures. It was a similar story among provincial governments, which provide most of the operating support to universities. Davidson says that they were forced to reduce their faculty ranks by about 4,000 to around 31,000 during the 1990s.

And all this when researchers in the United States seemed to be having it so good. There is still some debate as to the extent of Canada's brain drain. But Victor Ling, vice-president of research at the BC Cancer Agency in Vancou-



Strangway: building for the future.

ver, believes that "the quality of the [brain] drain is more important than the quantity".

Since putting its fiscal house in order, the federal government has set up the CFI and so far pumped almost Can\$2 billion (US\$1.3 billion) into the agency. It has restored funding for the granting councils to previous levels and renewed its commitment to the Centres of Excellence Programme. Moreover, it has formed the CIHR, established a national genome initiative (Genome

Canada catalyses health research with thirteen virtual institutes

One development energizing the research community in Canada at the moment is the arrival of the Canadian Institutes of Health Research (CIHR), which opened for business in June and provides the much-anticipated new structure of support for health research.

The most tangible difference between the CIHR and the Medical Research Council of Canada, which it replaces, is the level of federal government support. In the transition that created the CIHR, its budget was increased about 25% to Can\$402 million (US\$270 million) for the current year, and will rise by a further 33% to Can\$533 million next year.

"We don't have any further commitment from the

government but I'm hoping that they will see this as a phase one," says CIHR president Alan Bernstein. The omens are good. Bernstein says that as recently as last month, federal health minister Allan Rock said that he viewed the present figure not as a stopping point but as a start towards a Can\$1 billion budget for the CIHR.

A budget of this size would elevate the federal government's expenditure on health research to about 1% of its total healthcare spend, making it comparable to levels in the United Kingdom, for example. "I think the vision for CIHR when it's fully functional will require a budget of that order," Bernstein says.

Another change is that researchers with common research interests and goals, but spread about in universities, hospitals and research centres across the country, will now be linked together in 13 'virtual' institutes (see list, overleaf), the final selection of which was made in July (see *Nature* 406, 445; 2000).

Many have been quick to draw parallels between the CIHR and the US National Institutes of Health (NIH) and some of the CIHR's institutes do bear a striking resemblance to those in the NIH's intramural programme. But unlike the NIH, the CIHR will not have an intramural component. Bernstein says that by concentrating the CIHR's resources in universities and

Canada), and begun to address the human resource issue.

In terms of the Can\$160 million Genome Canada initiative, there is a sense that "we missed the boat a little bit in the first wave of genome research", says Christian Sylvain, senior policy analyst at the AUCC.

Nevertheless, Genome Canada will establish five genome centres, or hubs, probably in Halifax, Montreal, the Prairies, Toronto and Vancouver, with the aim of developing the kind of critical mass needed to carry out competitive research, as well as establishing core facilities for use by researchers across Canada.

The human-resource issue in Canada is not just about making up the shortfall created in the 1990s. Davidson says that the ageing of Canada's faculty is also a concern, with the average age now about 49. Added to this, the AUCC estimates that student enrolment will increase by 20% over the next decade, so universities will need to hire 2,500–3,000 new faculty each year over this period.

This might be a tall order given that Canada produces about 4,000 PhDs a year and only some of these will end up in academia. Universities will need to increase their share of these PhD graduates, says Davidson, as well as attract individuals from outside Canada. Programmes such as the CFI's New Opportunities Fund, and the 21st Century Chairs for Research Excellence programme, which will create 2,000 new research chairs over the next five years, will certainly help.

Faculty renewal is going hand in hand with institutional renewal. With all the new initiatives to kick-start careers and recruit and retain leading scientists, the "biggest limiting factor at the moment is space", says Jim Woodgett of the medical biophysics department at the Ontario Cancer Institute in Toronto, not so much in terms of availability of funds but because of the time it takes to build.

The CFI's Innovation Fund, was created with infrastructure development in mind. Although it is not intended to address the issue of deferred maintenance, the fund does provide support for equipment, buildings



Helping hand: the BC Cancer Agency has received a major grant to build a new cancer research centre, replacing its existing facility which is housed in a converted bakery.

and core facilities. The CFI contributes 40% of 'eligible' project costs, which is usually matched by provincial governments, leaving the institution to find the rest.

In the most recent round of awards, announced in July, one of the major beneficiaries was the BC Cancer Agency, which will receive Can\$27.8 million from the CFI towards a new Can\$82 million, 225,000-square-foot cancer research centre. When built, it will provide a fully integrated cancer research facility that will also house the Genome Sequence Centre headed by Nobel laureate Michael Smith.

The new cancer research building is part of a joint initiative between the BC Cancer Agency and the University of British Columbia (UBC) to form a Centre for Integrated Genomics headed jointly by Ling and Douglas Kilburn, director of the UBC's Biotechnology Laboratory. The centre will encompass a new Can\$30 million biotechnology laboratory being planned at the UBC, for which it received Can\$9.3 million from the CFI last year, as well as the new cancer research centre.

Provincial governments are also investing more aggressively in research and development, and encouraging greater collaboration between the public and private sectors in an effort to boost regional economies.

In Alberta, for example, which is prospering as Canada's main oil-producing region, the provincial government announced in February a Can\$500 million endowment fund to support R&D in basic science and engineering, which will be administered by the new Alberta Heritage Foundation for Science and Engineering Research. And in April the Ontario Research and Development Challenge Fund announced plans to invest up to Can\$75 million over five years in genomics research in the region.

This current transformation in Canada's research enterprise is a consequence of no one initiative but rather "the congruence of all of them", says Woodgett. Although the resources are somewhat focused on areas that have the highest potential, the important point is they are "not being diluted, which has always been the Canadian way", he says. ■

teaching hospitals, "as well as funding health research, we're strengthening the education system and healthcare system".

The CIHR is in the midst of a national search for institute directors and advisory board members. Nominations are flooding in and Bernstein hopes to announce the names of the institute directors in the next six weeks. "The actual institutes are perhaps less important than the philosophy that goes into how they will work and the quality of the individuals that will lead and advise them," Bernstein says.

Once in place, the directors (who will remain at their home institutions), together with their institute advisory boards, will be asked to develop a strategic plan for their institute. "For the first time we'll have at the national level an organization which is not simply

writing out cheques but actually setting the health research agenda for Canada," says Bernstein. It will put health research where it should be, he says, "which is front and centre".

Although the CIHR received a healthy budget increase, its mandate has been broadened to make it all-inclusive. The four pillars of health research — biomedical, clinical, population and health services — now all come under the CIHR umbrella. Bernstein says that he has already increased the number of grant panels in the investigator-initiated area to reflect this broader mandate.

Research training is also under the CIHR's purview and Bernstein plans to look into the possibility of providing institutional training grants, which have been used to great effect in the United

States. Under the current system it is individuals who apply. ■

Diane Gershon is Assistant Editor, New Technology, at Nature Medicine.

♦ <http://www.cihr.ca>

The CIHR's virtual institutes are: Aboriginal Peoples' Health; Cancer Research; Circulatory and Respiratory Health; Gender and Health; Genetics; Health Services and Policy Research; Healthy Aging; Human Development, Child and Youth Health; Infection and Immunity; Neurosciences, Mental Health and Addiction; Musculoskeletal Health and Arthritis; Nutrition, Metabolism and Diabetes; and Population and Public Health.